An Overview of Signal Processing Techniques for Terahertz Communications

Hadi Sarieddeen, Member, IEEE, Mohamed-Slim Alouini, Fellow, IEEE, and Tareq Y. Al-Naffouri, Senior Member, IEEE

Abstract—Terahertz (THz)-band communications are a key enabler for future-generation wireless communication systems that promise to integrate a wide range of data-demanding applications. Recent advancements in photonic, electronic, and plasmonic technologies are closing the gap in THz transceiver design. Consequently, prospect THz signal generation, modulation, and radiation methods are converging, and the corresponding channel model, noise, and hardware-impairment notions are emerging. Such progress paves the way for well-grounded research into THz-specific signal processing techniques for wireless communications. This tutorial overview these techniques with an emphasis on ultra-massive multiple-input multiple-output (UMIMO) systems and reconfigurable intelligent surfaces, which are vital to overcoming the distance problem at very high frequencies. We focus on the classical problems of waveform design and modulation, beamforming and precoding, index modulation, channel estimation, channel coding, and data detection. We also motivate signal processing techniques for THz sensing and localization.

Index Terms—THz communications, signal processing, ultra-massive MIMO, intelligent reflecting surfaces, terabit per second.

I. INTRODUCTION

The wireless communication’s frequency spectrum has been continuously expanding in an attempt to satisfy the ever-increasing bandwidth demands. While millimeter-wave (mmWave)-band communications [1], [2] are already shaping the fifth-generation (5G) of wireless mobile communications, terahertz (THz)-band communications [3]–[8] are expected to play an essential role in the future sixth-generation (6G) [9]–[20] and beyond. As such, THz-related research has attracted significant funding, and standardization efforts have been launched [21]–[23]. Being the last unexplored piece of the radio-frequency (RF) spectrum, the THz band is sandwiched between the microwave and optical bands. Hence, technologies from both sides are being explored to support THz communications. RF engineers label as THz all operations beyond the 100 GHz threshold, below which most known mmWave use cases exist. Optical engineers, on the other hand, label as THz any frequency below 10 THz (the far-infrared). But, the THz range is 300 GHz–10 THz according to IEEE Transactions on Terahertz Science and Technology, and closely mapped to the Tremendously High Frequency (THF) band (300 GHz–3 THz) according to ITU-R.

THz communications are called-for despite the maturity of neighboring technologies. Unlike mmWave communications, THz communications can exploit the available spectrum to achieve a terabit/second (Tbps) data rate without additional spectral efficiency enhancement techniques. Furthermore, due to the shorter wavelengths, THz systems can support higher link directionality, are less susceptible to free-space diffraction and inter-antenna interference, can be realized in much smaller footprints, and possess higher resilience to eavesdropping. On the other hand, compared to visible light communications (VLC) [24], [25], THz signals are not severely affected by alignment issues, ambient light, atmospheric turbulence, scintillation, fog, and temporary spatial variation of light intensity. THz communications can thus complement both mmWave and VLC by providing alternative quasi-optical paths. However, due to significant water vapor absorption above 1 THz, naturally, a gap might always exist for wireless communications at the high end of the THz range.

THz communications are thus expected to enable ultra-high bandwidth and ultra-low latency communication paradigms [26]. For example, they can be used to achieve optical-fiber-like performance in network backhauling [16], backbone (rack-to-rack) connectivity in data centers [27]–[29], and high data rate kiosk to mobile communications [30]. Furthermore, THz wireless bridges enable transparent integration of fiber networks without requiring detection, decoding, and re-modulation [31]. THz links supporting 100 gigabit/second (Gbps) have already been demonstrated over distances corresponding to such applications [32]. However, the holy grail of THz communications is enabling mobile communications at the device level and the access level in the context of medium-range indoor, vehicular, drone-to-drone, or device-to-device communications. Specifically, when combined with other THz-band applications such as accurate localization, sensing, and imaging, THz communications can enable wireless remoting of human cognition, leading to ubiquitous wireless intelligence [26], [33].

The main contributions to THz technology are still at the device level rather than the system level. High-frequency electromagnetic radiation is perceived either as waves that get treated via electronic devices (the mmWave realm) or as particles that get processed via photonic devices (the optical realm). In between, the THz band is dubbed as a “THz gap” due to the lack of compact THz signal sources and detectors that have high power and sensitivity, respectively. Recent advancements

The authors are with the Department of Computer, Electrical and Mathematical Sciences and Engineering (CEMSE), King Abdullah University of Science and Technology (KAUST), Thuwal, Makkah Province, Kingdom of Saudi Arabia, 23955-6900 (e-mail: hadi.sarieddeen@kaust.edu.sa; slim.alouini@kaust.edu.sa; tareq.alnaffouri@kaust.edu.sa). This work was supported by the KAUST Office of Sponsored Research.
in electronic and photonic THz transceiver design, however, allowed efficient signal generation, modulation, and radiation \cite{32,34-36}. Electronic solutions, which are mainly based on silicon complementary metal-oxide-semiconductor (CMOS) and silicon-germanium BiCMOS technologies \cite{37-39}, demonstrated incredible compactness and compatibility with existing fabrication processes. However, the corresponding highest unity current gain frequency ($f_T$) and unity maximum available power gain frequency ($f_{\text{max}}$) remain at 280 gigahertz (GHz) and 320 GHz, respectively. Nevertheless, higher operating frequencies have been noted by III-V-based semiconductors \cite{40} in high electron mobility transistors (HEMTs) \cite{41,42}, heterojunction bipolar transistors (HBTs) \cite{43}, and Schottky diodes \cite{44}. In photonic devices \cite{45}, where the main design driver is the data rate, higher carrier frequencies are supported, but the degrees of integration and output power are low. Frequencies beyond 300 GHz have been supported using optical downconversion systems \cite{32}, quantum cascade lasers \cite{45}, photoconductive antennas \cite{46}, and uni-traveling carrier photodiodes \cite{47}.

Satisfying emerging system-level properties requires designing efficient and programmable devices. This deviation from designing perfect THz devices resulted in the emergence of integrated hybrid electronic-photonic systems \cite{35}. What also gaining popularity is plasmonic solutions \cite{48,49} where novel plasmonic materials such as graphene possess high electron mobility and reconfigurability \cite{50-53}. The resultant surface plasmon polariton (SPP) waves in plasmonic antennas have much smaller resonant wavelengths than free space waves, which results in compact and flexible designs. By leveraging the properties of plasmonic nanomaterials and nanostructures, transceivers and antennas that intrinsically operate at THz frequencies can be created, which avoids the upconversion and downconversion losses of electronic and photonic systems, respectively. Graphene can be utilized to develop direct THz signal sources, modulators (that manipulate amplitude, frequency, and phase) and on-chip THz antenna arrays \cite{54}. All these advances show that the gaps germane to designing THz technology are rapidly closing and that the THz-band will soon open for every-day applications.

Yet many challenges still need to be addressed from a signal processing and communication system’s perspective. The factors to be considered in signal processing in the THz realm differ greatly from those in the systems at lower frequencies; they are closely linked to the transceiver or device architectures. Efficient THz-band signal processing is crucial for two reasons. First, it must account for the use ultra-massive multiple input multiple output (UM-MIMO) antenna systems \cite{55-57} to overcome the very short communication distances due to severe power limitations and propagation losses. Second, it must overcome the mismatch between the bandwidth of the THz channel and that of the digital baseband system \cite{58,59}. Since channel coding is the most computationally demanding component of the baseband chain, several projects are studying efficient coding schemes for Tbps operations \cite{60}. Moreover, the complete chain should be efficient and parallelizable. Therefore, joint algorithm and architecture co-optimization of channel estimation, channel coding, and data detection is required. Also, the inherent sparsity at THz can be exploited in solutions based on compressed sensing techniques. Furthermore, low-resolution digital-to-analog conversion systems can reduce the baseband complexity; all-analog THz solutions are even considered.

Although THz communications possess quasi-optical traits, they retain several microwave characteristics. They can still use UM-MIMO antenna array processing techniques to support efficient beamforming and reflective surfaces to support non-line-of-sight (NLoS) propagation. Efficient beamforming and beamsteering techniques, as well as low-complexity precoding and combining algorithms, are thus required. Yet, what seems predicted to be the norm in future THz systems are hybrid and adaptive arrays-of-subarrays (AoSA) antenna architectures, in which each subarray (SA) undergoes independent beamforming. Furthermore, given the large degrees of freedom in THz UM-MIMO systems at the transmitter side, a variety of probabilistic shaping and index modulation schemes can be explored. This is particularly true in plasmonic solutions where each AE can be turned on and off or assigned a specific frequency by simple material doping or electrostatic bias.

Molecular absorptions, further, result in band splitting and spectrum shrinking at larger communication distances. Distance-adaptive solutions in which antenna array designs and resource allocation criteria are optimized are required to tackle spectrum shrinking \cite{61,62}. Towards this end, the classical problems of waveform design and modulation need to be revisited. For instance, single-carrier (SC) modulations can be favored over orthogonal frequency-division multiplexing (OFDM), which is very complex to implement in the THz band. Nevertheless, in some indoor THz scenarios, a few multipath components might persist, which would result in frequency-selective channels. Frequency-selectivity might arise at the receiver side also due to the behavior of THz components. Therefore, multi-carrier modulations might still be required, perhaps in the form of multiple orthogonal non-overlapping SCs with some kind of carrier aggregation.

In this tutorial, we provide an overview on recent advancements in signal processing techniques for THz communications. We summarize and formulate the problem definitions of several THz-specific signal processing problems, and we further motivate open research problems. The paper is organized as follows: The system model is first presented in Sec. \[I\] followed by a discussion on THz channel and noise modeling in Sec. \[II\]. Then, recent performance analysis frameworks and experimental testbeds are summarized in Sec. \[IV\]. After that, the latest advances in THz modulation and waveform designs are summarized in Sec. \[V\] the concept of THz “spatial tuning” is proposed in Sec. \[VI\] and THz beamforming and precoding techniques are discussed in Sec. \[VII\]. The baseband signal processing problems of THz channel estimation, channel coding, and data detection are then illustrated in Sec. \[VIII\]. Afterward, an extension is made to treat the signal processing aspects of intelligent reflecting surface (IRS)-assisted THz communications in Sec. \[IX\]. Finally, in Sec. \[X-A\] we shed light on the importance of signal processing for THz sensing, imaging, and localization, and briefly discuss THz networking and security, before concluding in Sec. \[XI\].
II. SYSTEM MODEL

It is challenging to define a generic system model for THz communications at this early stage. Nevertheless, the use of AoSAs of antenna elements (AEs) is most likely to be the norm in future THz systems, as dynamic array gains are crucial for combating the distance problem. A typical THz communications system model is illustrated in Fig. [1] where adaptive AoSAs are configured at the transmitting and receiving sides. After digital-to-analog conversion (DAC) and before analog-to-digital conversion (ADC), each SA is fed with a dedicated RF chain. Due to high directivity, each SA is effectively detached from its neighboring SAs in a multi-user setting, and the role of baseband precoding reduces to defining the utilization of SAs, or to simply turning SAs on and off. In a point-to-point setup, however, SA paths can be highly correlated due to low spatial resolution.

We hereby adopt the three-dimensional (3D) UM-MIMO model of [63]–[65]. Let bold upper case, bold lower case, and lower case letters correspond to matrices, vectors, and scalars, respectively, and let (·)\textsuperscript{T} and (·)\textsuperscript{H} stand for transpose and conjugate transpose, respectively. The AoSAs consist of \( M_t \times N_t \) and \( M_r \times N_r \) SAs, at the transmitter and the receiver, respectively. Each SA is composed of \( Q \times Q \) AEs. Therefore, the overall configuration can be represented as a “large” \( M_t N_t Q^2 \times M_r N_r Q^2 \) MIMO system [66]. Such large or symmetric doubly-massive MIMO systems [67] are different from conventional massive MIMO systems. In the latter, large antenna arrays are typically configured at a transmitting base station to serve multiple single-antenna users at the receiver. The distances separating two SAs or two AEs are crucial for combating the distance problem. A typical LoS transmission over an SC frequency-flat fading channel.

Consequently, we assume for the generic system model an LoS transmission over an SC frequency-flat fading channel. The corresponding baseband system model is

\[
y = W_r^H H W_t^H x + W_n^H n, \tag{1}
\]

where \( x = [x_1 x_2 \cdots x_{N_t}]^T \in \mathbb{C}^{N_r \times 1} \) is the information-bearing symbol vector of components belonging to a constellation \( \mathcal{X} \) of quadrature amplitude modulation (QAM), for example, \( y \in \mathbb{C}^{N_r \times 1} \) is the received symbol vector, \( H = [h_1 h_2 \cdots h_{M_t N_r}] \in \mathbb{C}^{M_r N_r \times M_t N_t} \) is the channel matrix, \( W_t \in \mathbb{C}^{N_t \times M_t N_t} \) and \( W_r \in \mathbb{C}^{N_r \times M_r N_r} \) are the baseband precoder and combiner matrices, and \( n \in \mathbb{C}^{M_r N_r \times 1} \) is the additive white Gaussian noise (AWGN) vector of power \( \sigma^2 \).

An element of \( H, h_{m,n,m',n'} \), the frequency response between the \((m_t, n_t)\) and \((m_r, n_r)\) SAs, is thus defined as

\[
h_{m,n,m',n'} = a_t^H(\phi_t, \theta_t) G_r \alpha_{m,n,m',n'} G_t a_r(\phi_r, \theta_r), \tag{2}
\]

for \( m_t = 1, \ldots, M_t, \ n_t = 1, \ldots, N_t, \ m_r = 1, \ldots, M_r, \) and \( n_r = 1, \ldots, N_r \), where \( \alpha \) is the path gain, \( a_t \) and \( a_r \) are the transmit and receive SA steering vectors, \( G_t \) and \( G_r \) are the transmit and receive antenna gains, and \( \phi_t, \theta_t \) and \( \phi_r, \theta_r \) are the transmit and receive angles of departure and arrival, respectively (\( \phi \)'s are the azimuth angles and \( \theta \)'s the elevation angles). The steering vectors can be expressed as a function of the transmit and receive mutual coupling matrices, \( C_t, C_r \in \mathbb{R}^{Q^2 \times Q^2} \) as \( a_t(\phi_t, \theta_t) = C_t a_0(\phi_t, \theta_t) \) and \( a_r(\phi_r, \theta_r) = C_r a_0(\phi_r, \theta_r) \). By setting \( C_r = C_t = I_{Q^2} \) (identity matrix of size \( N_r \) ), the effect of mutual coupling is neglected. Such an assumption is valid in the plasmonic case, for example, for values of \( \delta_t, \delta_r \geq \lambda_{\text{app}} \) [68], where the SPP wavelength, \( \lambda_{\text{app}} \), is much smaller than the free-space wavelength, \( \lambda \). The ideal SA steering vector at the transmitter side can thus be expressed as

\[
a_0(\phi_t, \theta_t) = 1 \bigg/ \sqrt{Q} [e^{j\Phi_0,1,1}, e^{j\Phi_0,1,2}, \cdots, e^{j\Phi_0,1,Q}, \cdots, e^{j\Phi_0,1,1}, \cdots, e^{j\Phi_0,1,1} ]^T,
\]

where \( \Phi_{p,q} \) is the phase shift that corresponds to AE \((p, q)\), and is defined as

\[
\Phi_{p,q} = \psi_{x}^{(p,q)} \frac{2\pi}{\lambda_{\text{app}}} \cos \psi_{x}^{(p,q)} \sin \theta_t + \psi_{y}^{(p,q)} \frac{2\pi}{\lambda_{\text{app}}} \sin \psi_{x}^{(p,q)} \sin \theta_t + \psi_{z}^{(p,q)} \frac{2\pi}{\lambda_{\text{app}}} \cos \psi_{x}^{(p,q)} \cos \theta_t,
\]

with \( \psi_{x}^{(p,q)}, \psi_{y}^{(p,q)}, \) and \( \psi_{z}^{(p,q)} \) being the coordinate positions of AEs in the 3D space. At the receiver side, \( a_0(\phi_r, \theta_r) \) can be similarly defined.
III. THZ-BAND CHANNEL MODELING

Channel modeling is essential for efficient signal processing in the THz band. Accurate THz channel models should consider the effect of both the spreading loss and the molecular absorption loss and should account for the LoS, NLoS, reflected, scattered, and diffraacted signals. Channel modeling approaches are mainly deterministic or statistical [69]. While deterministic channel modeling uses computationally extensive ray-tracing techniques to capture site geometry, matrix-based statistical modeling represents each independent sub-channel by a random variable of a specific distribution. Hybrid channel modeling schemes combine the advantages of both approaches, where dominant paths are captured deterministically, and other paths are statistically generated.

A. Ray-Based THz Channel Modeling

Several extensive ray-tracing-based THz propagation measurements have been recently reported. For instance, a unified multi-ray THz-band channel model is proposed in [70], which covers the LoS, scattered, reflected, and diffracted paths, and which is experimentally validated over 0.06 – 1 THz. In [71], a deterministic channel model for 0.1 – 1 THz is proposed for LoS and NLoS scenarios, using the Kirchhoff scattering theory and ray tracing. Similarly, ray-based sub-THz channel characterization at 90 – 200 GHz is detailed in [72] using deterministic simulations in indoor office and outdoor in-street scenarios. Other THz ray-tracing channel modeling attempts are tailored for the peculiarities of specific use cases. For instance, a ray-tracing channel model at 300 GHz is presented in [30], for close-proximity THz communications such as in the case of Kiosk downloading. Also, at 300 GHz, a ray-tracing simulator with calibrated electromagnetic parameters is used in [73] for vehicle-to-infrastructure THz communications. Note that to reduce the complexity of ray tracing in UMi-MIMO systems, select few virtual paths can be captured, say between virtual transmitting and receiving points, the response of which gets mapped to actual pairs of transmitting and receiving AEs.

B. Statistical THz Channel Modeling

As an alternative to time-consuming and complex ray-tracing models in fixed geometries, several statistical THz channel modeling attempts are noted. For example, by developing a wideband channel sounder system at 140 GHz, indoor wideband propagation and penetration measurements for common building material are reported in [74]. In [75], indoor measurements and models for reflection, scattering, transmission, and large-scale path loss are also provided by the same group, for mmWave and sub-THz frequencies. Lower reflection loss is noted at higher frequencies in indoor drywall scenarios (stronger reflections). Partition loss, on the other hand, increases due to more prominent depolarizing effects. In [76], the statistical characterization of three channel bands between 300 THz and 400 THz is presented based on a broad set of measurements in LoS and NLoS environments and including spatial and temporal variations. The large-scale losses are modeled using the single slope path loss model with shadowing, where variations due to shadowing are shown to be normally distributed. Metal, wood, and acoustic ceiling panels prove to be good reflectors. In indoor environments, strong multipath components are shown to reduce the coherence bandwidth significantly, and high channel correlation is noted. Using a virtual antenna array technique, the same testbed was exploited in [77] to demonstrate 2 x 2 THz LoS MIMO channels.

Another stochastic indoor 300 GHz spatio-temporal channel model is introduced in [78], which takes into account parameters such as polarization, ray amplitudes, times of arrivals, angles of arrival and departures, and path-specific frequency dispersion. Furthermore, THz channel modeling via a mixture of gamma distributions is proposed in [79]. In other notable works, LoS broadband THz channel measurements are reported in [80] when using convex lenses at the transmitter, that are coupled with collimating lenses at the receiver. A beam domain channel model is also introduced in [81]. With large numbers of base stations and users, and given that several wavelengths typically separate users at high frequencies, it is shown that the beam-domain channel elements are statistically uncorrelated [82], and that their envelopes are independent of frequency and time. Moreover, a geometric-based stochastic time-varying model at 110 GHz is proposed in [83] for THz vehicle-to-infrastructure communications.

Statistical modeling can also be used to study the effect of blockage, which is significant at higher frequencies. For instance, NLoS THz channel modeling is conducted in a generic stochastic approach in [84], by assuming rectangular geometry and accounting for variable densities of reflecting objects (single reflection components) and blocking obstacles. THz signals are more sensitive to blockages than mmWave signals. Human blockage in indoor THz communications is studied in [85], where adding extra antennas to account for blocked streams is considered. In [86], the dynamic blockages caused by moving humans are incorporated into the 3D THz channel model, since deploying antennas at high altitudes would help in preventing blockages. In fact, blockages can arise at the transmitter due to the condensation of particles. The latter can be mitigated by compressed sensing techniques [87], [88].

C. The Molecular Absorption Effect

The path loss seen by a THz signal in the presence of water vapor is dominated by spikes that represent molecular absorption losses originating at specific resonant frequencies, due to excited molecule vibrations. Higher densities of absorbing molecules make the peaks stronger and wider (broadening of absorption lines). Because of these lines, the spectrum gets divided into smaller windows (sub-bands), each of which has a width of tens or hundreds of GHz. These windows are distance-dependent since some spikes only get significant at specific distances (by increasing the distance from 1 to 10 meters, the transmission windows are reduced by order of magnitude [89]). Hence, variations in the communication distance affect both the available bandwidths and the path loss (the available bandwidth shrinks at higher frequencies).
The LoS path gain as a function of absorption is expressed as
\[
\alpha_{LoS}^{m_r,n_r,m_n,n_t} = \frac{c}{4\pi f d_{m_r,n_r,m_n,n_t}} \times e^{-\frac{1}{2}K(f)d_{m_r,n_r,m_n,n_t}} e^{-\frac{1}{2}f d_{m_r,n_r,m_n,n_t}},
\]
where \( d_{m_r,n_r,m_n,n_t} \) is the distance between the transmitting and receiving SAs, \( K(f) \) is the absorption coefficient, \( f \) is the frequency of operation, and \( c \) is the speed of light in vacuum. \( K(f) \) is derived in [48] as a summation over contributions from isotopes \( i \in \{1, \ldots, I\} \) of gases \( g \in \{1, \ldots, G\} \) that constitute a medium. The construction in [48] uses radiative transfer theory to give insight on the physical meaning of the corresponding equations, as a function of temperature, system pressure, and absorption cross section. We hereby compile the overall equation for \( K(f) \) in [3], where \( T \) is the system temperature, \( K_B \) is the Boltzmann constant, \( p \) is the system pressure, \( q^i,g \) is the mixing ratio of gas \( i \), \( f^{i,g} \) is the resonant frequency at reference pressure, \( \gamma \) is the temperature broadening coefficient, \( \delta^{i,g} \) is the line pressure shift of gas \( i \), \( S^{i,g} \) is the line intensity, \( \alpha_0^{air} \) is the broadening coefficient of air, and \( \alpha_0^{g} \) is the broadening coefficient of gas \( i \). Note that all these parameters can be extracted from the high-resolution transmission molecular absorption database (HITRAN) [90]. However, this model is complex and hard to track analytically.

Since the effect of water vapor dominates the absorption losses at high frequencies, a simplified yet sufficiently accurate model for molecular absorption loss is developed in [91], [92] and used in [93], [94]. This model is built in a database approach by fitting the absorption line shape functions to the actual responses. It is mainly tailored for the 0.1-0.45 THz band, but it also applies for specific sub-bands within the range. In particular, the absorption coefficient \( K(f) \) can be approximated as
\[
K(f) = K_1(f,v) + K_2(f,v) + K_3(f),
\]
where
\[
K_1(f,v) = \frac{A(v)}{B(v)} \left( \frac{f}{100c} - c_1 \right)^2
\]
\[
K_2(f,v) = \frac{C(v)}{D(v)} \left( \frac{f}{100c} - c_2 \right)^2
\]
\[
K_3(f) = \rho_1 f^3 + \rho_2 f^2 + \rho_3 f + \rho_4,
\]
with \( v \) being the volume mixing ratio of the water vapor. The coefficients \( \rho_1, \rho_2, \rho_3 \) and \( \rho_4 \) and the functions \( A,B,C \) and \( D \) are detailed in [92]. Nevertheless, the exact HITRAN-based absorption model is still favored, especially in very high signal-to-noise ratio (SNR) settings, and in the context of joint signal processing for communications and sensing (Sec. X-A).

D. Effect of Scattering

At higher frequencies, the electromagnetic roughness of surfaces increases, which causes diffuse scattering and increased backscattering (at lower incident angles) [96]. The effect of scattering on the reflection coefficient in THz communications is studied in [97]. It is shown that the scattered power increases with frequency and surface roughness relative to the reflected power, where smooth surfaces (like drywall) can be modeled as reflective surfaces. In [98], diffuse scattering in THz massive MIMO channels is studied by developing a hybrid modeling approach for 3D ray-tracing simulations, assuming realistic indoor environments over the 0.3-0.35 THz band. The channel capacity of indoor massive MIMO channels is calculated, assuming different surface roughnesses for LoS and NLoS scenarios. It is shown that scattering can be exploited to enhance the spatial multiplexing gains. Note also that diffuse scattering can be exploited to identify the type of surfaces.

In [99], [100], it is argued that THz beams are more susceptible to snow than rain, suffering higher losses under an identical fall rate Mie theory approach for electromagnetic radiation. THz-band rain-induced co-channel interference is studied in [101], using the bistatic radar and the Mie scattering theory, and assuming first-order multiple scattering. It is illustrated that the overall interference levels due to rain are significantly lower in the THz band (20 dB difference between 300 GHz and 60 GHz), except when the receiver is very close to the LoS (forward-oriented scattering at high frequencies).

E. Effect of Misalignment and Impairments

The performance of THz communications systems is severely deteriorated under the effect of misalignment and hardware impairments. The joint effect of misalignment and hardware impairments on THz communications is studied in [102], including in-phase and quadrature imbalance (IQI) and non-linearities. The study models all impairments as Gaussian noise components, and accounts for operation and design parameters, alongside environmental parameters; it also introduces a misalignment fading model, which is of crucial importance for THz communications. The work in [102] is extended in [103] to capture the error analysis of mixed THz-RF wireless systems. In fact, the use of high-directivity antennas in THz systems results in small transceiver antenna beamwidths, which, although they provide higher antenna gains, they cause pointing errors and loss of connection. Due to the symmetry of the beam, the misalignment fading component depends only on the radial distance [104]. Earlier attempts
to capture this misalignment effect are reported in [105]–[107]. Furthermore, the effect of small-scale mobility on THz systems is studied in [108], [109]. It is shown that simple shakes or rotations due to user equipment mobility can result in beam misalignment and SNR degradation, which in turn result in a loss in communication time due to extra beam search mechanisms. Therefore, there exists a trade-off between antenna directivity and capacity.

Misalignment can be modeled by expressing the effective channel coefficient between two SAs in terms of three components as $h_{\text{eff}} = h h_{\text{ma}} h_d$, where $h$ is expressed in [2]. $h_{\text{ma}}$ represents the misalignment fading, and $h_d$ is the stochastic path gain (can be neglected or modeled as an $\alpha - \mu$ process depending on the scenario). Due to beam symmetry, misalignment fading mainly depends on the pointing error, which can be expressed in the form of a radial distance $r$ between the transmission and reception beams at a communication distance $d$:

$$h_{\text{ma}}(r; d) \approx A_0 \exp\left(-\frac{2r^2}{w_{\text{eq}}^2}\right),$$

where $A_0$ is the fraction of power collected at the receiver, and $w_{\text{eq}}$ is the equivalent beamwidth. As for hardware imperfections in both the transmitter and the receiver, they can be modeled as two additional distortion noises. The modified system model under impairments is approximated as

$$y = H(x + n_t) + n_r + n,$$

where $n_t \in \mathbb{C}^{M_t \times N_t \times 1}$ and $n_r \in \mathbb{C}^{M_r \times N_r \times 1}$ are two complex Gaussian distortion noise vectors at the transmitter and the receiver, with noise variances $\eta_t^2 \hat{P}$ and $\eta_r^2 \hat{P}|h|^2$, respectively, with $\hat{P}$ being the average transmitted power and $\eta_t$ and $\eta_r$ being the impairment coefficients [102].

### F. Multipath THz Channels

Despite only considering an LoS-dominant scenario in our system model, a multipath channel can arise in several THz communications scenarios, especially indoors, where lower antenna gains can be tolerated. Nevertheless, it is safe to assume that a THz channel is sparser than a mmWave channel. For instance, only 5 multipath components survive at 0.3 THz in a 256 $\times$ 256 UM-MIMO system, which is 32.5% less than the number of multipath components in the same system at 60 GHz [110]. On average, the path gain difference between LoS and NLoS paths is 15 dB higher in THz systems compared to that in mmWave systems [63]. The angular spread of indoor THz channels, however, is much smaller than that at lower frequencies [65]. Indoor THz multipath components [78], [111], [112] can be modeled using the Saleh-Valenzuela (S-V) channel model [113], where the channel response within a time margin $T_s$ is expressed as

$$h_{\text{NL,S}}^{\text{NL,S}}(f, d) = \sum_{i=0}^{N_{\text{LoS}} - 1} \sum_{l=0}^{N_{\text{Ma}}} \alpha_{il}(f, d) G_r\left(\phi_{il}, \theta_{il}\right) G_t\left(\phi_{il}, \theta_{il}\right) \times a_r\left(\phi_{il}, \theta_{il}\right) a_t\left(\phi_{il}, \theta_{il}\right),$$

where $N_{\text{eq}}^{(i)}$ and $N_{\text{clus}}$ are the number of rays in the $i$th cluster and the number clusters, respectively. The path gain coefficient, $\alpha_{il}$, can be calculated as detailed in [114].

### G. Ultra-Wideband THz Channels

Ultra-wideband THz channels can arise in several scenarios, especially with pulse-based modulations, which are serious candidates for THz communications. Such short pulses in time span the entire THz range in frequency (more on that in Sec. V). The corresponding time-domain channel responses are captured in [115], [116] for short-range (less than 1 m) true THz communications (0.1 – 10 THz), where the effects of both molecular absorption and rough surfaces is studied.

At THz frequencies, the receiving antenna array might often be larger than the received beamwidth, which means that a path is not visible to all the antennas in the array. This results in non-stationarity [117] and variations in the time of arrival, angle of arrival (as well as the angle of departure), and receive amplitudes across the antenna array. This problem is more serious when transmitting THz signals with large bandwidths. In particular, the variation of the time of arrival would cause inter-symbol interference (ISI) in wideband scenarios (instead of only a phase shift). Such spatial and frequency dual-wideband effects are captured in [118] in the context of mmWaves, where the channel sparsity is exploited in both the angle and delay domains.

Furthermore, a deterministic frequency-selective fading, which can not be captured by Rayleigh and Ricean models, arises due to molecular absorption. This fading results in delayed signal components, especially in a wideband multipath scenario. Hence, group velocity dispersion (GVD) is another issue that arises in impulse radio THz communication due to frequency-dependent refractivity in the atmosphere. GVD becomes limiting at specific link distances, atmospheric water vapor densities, and channel bandwidths [119]. This phenomenon also results in ISI as data bits spread out of their assigned slots and interfere with neighboring slots. Therefore, dilating bit slots can solve this problem at the expense of data rate. In [120], the atmospheric GVD of THz pulses (0.2 – 0.3 THz) is compensated using stratified media reflectors.

### H. THz Noise Modeling

Accurate noise models are important for understanding the behavior of THz systems. Despite the fact that stochastic models for the electronic noise at THz receivers are still lacking, in [121], two primary sources of noise are noted: thermal noise, which arises at the receiver multiplier and mixer chains, and absorption noise that is channel induced due to water vapor molecules. The corresponding histogram of the measured noise is shown to follow a Gaussian distribution, which is in accordance with the noise behavior at lower frequencies, as opposed to shot noise in optical receivers. In [122], the transmission-induced noise due to molecular absorption is discussed in more detail. The authors differentiate between multiple models, most of which are based on the antenna temperature that is generated by the absorbed energy. However, this molecular absorption noise model has
never been validated by measurements; it uses sky noise as a basis, which might overestimate the level of the self-induced noise. In [123], molecular absorption is assumed to result in a rich scattering environment because absorptions are followed by re-radiations with minimal frequency shifts. Nevertheless, such coherent re-radiations can be more realistically lumped in a generic absorption noise factor [48, 122]. The channel-induced component thus dominates the overall noise in pulse-based systems (especially in low-noise graphene-based electronic devices [124]), and it is colored over frequency. The total noise power at a distance $d$ can be expressed as

$$\sigma^2 = K_B \int_B T_{\text{noise}}(f, d)df,$$

where $T_{\text{noise}} = T_{\text{sys}} + T_{\text{mol}} + T_{\text{other}}$, and where $T_{\text{sys}}$ and $T_{\text{mol}}$ are the system electronic noise temperature and the molecular absorption noise, respectively, and

$$T_{\text{mol}}(f, d) = T_0(1 - e^{-K(f)d}).$$

In a carrier-based system with perfect frequency planning over absorption-free spectra, the effect of the channel-induced noise can be minimized. This is particularly true at shorter communication distances, where propagation losses dominate. At larger distances and higher frequencies, however, molecular absorption losses can take over propagation losses. Note that an additional low-frequency noise component exists at the transmission chain and the power supply.

In addition to the above-mentioned sources of noise, the effect phase noise (PN) at high frequencies should be taken into consideration. Phase noise is caused by time-domain instability (jitter), which gives rise to random rapid, short-term fluctuations in the phase. Precise THz-specific PN measurements at THz are still lacking, despite some modeling attempts [125, 126]. PN is modeled by the superposition of Wiener and Gaussian noise. Furthermore, PN is typically accompanied by strong phase impairments and carrier frequency offset (CFO), where both result from poorly performing high-frequency oscillators (more of a problem in multi-carrier systems [127]). Novel signal processing and optimized modulation/demodulation can overcome these impairments and achieve PN robustness. THz CFO can be estimated via an under-sampling approach with narrow-band filtering and coprime sampling [128].

### IV. Analysis Frameworks and Testbeds

The more accurate the THz channel models get, the better the insight we have on the achievable gains of THz systems. These gains can be captured via theoretical performance analysis frameworks and can be verified via experimental testbeds. In the following, we summarize recent results on both.

#### A. Performance Analysis Frameworks

Early THz channel capacity studies are reported in [129] in the context of nanonetworks, where numerical results are generated for different molecular compositions and power allocation schemes, with emphasis on pulse-based modulation.

In [70], a thorough analysis of THz channel characteristics is presented, where the variability of spectral window widths over communication distances is first observed, and it is illustrated that water-filling power allocation can achieve more than 75 Gbps with 10 dBm transmit power over 0.06 – 1 THz. The study further shows that in multipath scenarios, the root-mean-square delay spread is distance- and frequency-dependent, and the coherence bandwidth is less than 5 GHz (decreases with longer distance and lower carrier frequencies). Furthermore, the spacings between transmissions are affected by the increased temporal broadening effects at higher frequencies, wider pulse bandwidths, and longer distances. Therefore, distance-adaptive techniques are suggested, alongside multi-carrier transmission schemes.

Theoretical performance analysis of conventional THz communication systems is also progressing. For instance, in [94], the two-path channel characteristics over 275 – 400 GHz are analyzed in terms of SNR and ergodic capacity, by taking into account channel characteristics such as frequency selectivity, path-loss, and atmospheric conditions. The analysis also accounts for transceiver parameters such as antenna gains and transmit power. By assuming the signal and noise to be jointly Gaussian, classic Shannon results for coherent reception are used for capacity estimation. This approach is employed in several other studies: It is used in [48] to compute the capacity of THz nano-sensor networks, in [130] to capture the relation between transmission distance and absorption-free (transparency) windows, and in [131] to study the data rates of fixed THz-links. Such studies are also extended in [53] in the context of reconfigurable MIMO systems. Furthermore, the effect of both deterministic (molecular absorption) and random (atmospheric turbulence and pointing errors) factors on the bit-error-rate performance and capacity of LoS THz links is studied in [132], where the log-normal, gamma-gamma and exponentiated Weibull channel models are employed.

THz systems are also very sensitive to transceiver imperfections. In [133], it is argued that in the presence of PN and misalignment, the outage performance is not significantly enhanced by higher transmit power, and lower-order modulations might be required. The effect of local oscillator hardware impairments, further, can be more severe than misalignment issues. The performance of THz systems is studied under the joint impact of PN and misalignment fading in [134], and under the joint impact of PN and amplifier non-linearities in [135]. The impact of these errors, when lumped into PN at the local oscillator, is further studied under different transceiver architectures in [136].

Stochastic geometry is also used for analyzing THz systems. For instance, a stochastic geometry approach for mean interference power and outage probability analysis is considered in [137], in the context of a dense THz network operating over 0.1 – 10 THz. The authors model the interference as a shot noise process, and they assume directional antennas. Compared to conventional applications, stochastic geometry techniques in THz networks should adapt to account for the inherent low fading and additional absorption losses. Furthermore, despite the fact that high antenna gains result in a lower probability of interference in the THz band, the interference...
level is much higher when it occurs. Stochastic geometry is also used in [138] to derive the exact and approximate distributions of the received signal power and interference, respectively; semi-closed-form expressions are derived for the coverage probability and the average achievable rate.

The achievable data rates of indoor THz systems are studied in [139], where a single frequency network is advocated, and the corresponding ISI due to channel dispersion is taken into consideration; the effect of the density of access points on performance is studied. Similarly, in [140], the indoor interference and coverage under beamforming are studied. An analytical model for the distribution of indoor access points at different blocks of the THz spectrum is then proposed in [141]. In THz indoor NLoS scenarios, the total received signal is mainly accumulated from diffuse reflections [142]. Small-scale mobility in indoor THz scenarios is also studied in [143] as a function of several variables, such as frequency windows, beamwidths, distance, humidity, mobility type, and antenna placement. It is observed that there exist optimal beamwidths for specific mobility types and AP placement strategies.

The performance of THz systems is also studied in use-case-specific scenarios. For example, the reliability and latency of THz communications are studied in the context of wireless virtual reality in [144], [145]. By deriving a tractable expression for system reliability as a function of THz system parameters, it is shown that with proper densification, high reliability can be achieved. In another example [146], the channel capacity and reliability are studied for the special case of THz wireless networks-on-chip communications [147], [148], where it is shown that performance can be enhanced by proper choice of silicon layers and their thickness. Furthermore, user- and network-centric metrics for THz information shower systems are evaluated in [149], where it is shown that 95% of traffic from long-range networks can be offloaded, initiating heavy-traffic THz information shower sessions.

THz signals are also being considered for communications at atmospheric altitudes (where the concentration of water vapor decreases) among drones, jets, unmanned aerial vehicles, as well as satellites. Calculating the absorption loss through the atmosphere at higher altitudes is decisive for enabling such applications. In [150], it is illustrated that communication at 0.75–10 THz is more feasible at higher altitudes than sea-level, with reported usable bandwidths of 8.218 THz, 9.142 THz, and 9.25 THz over a distance of 2 km. In [151], the capacity of optoelectronic THz Earth-satellite links is analyzed, where the claim is made that 10 Gbps per GHz can be supported. Similarly, the use of the THz band for simultaneously providing high data rates and wide coverage data streaming services to ground users from a set of hotspots mounted on flying drones is studied in [152]. By solving an optimization problem for resource utilization, it is shown that much higher throughput can be achieved in mobile environments compared to static environments. A holistic investigation on THz-assisted vertical heterogeneous networks, further, is conducted in [153]. The study comprises, in addition to terrestrial communication links, geostationary and low-earth orbit satellites, networked flying platforms, as well as in-vivo nano-networks; accurate channel modeling is key for harmony across all these applications.

B. Experimental Demonstrations

Several experimental demonstrations have been conducted to verify the corresponding channel models and the predicted performance metrics. Experimental results for the first true-THz absorption-defined window above 1 THz (1.02 THz) are reported in [121], where tens of Gbps are demonstrated in a multi-carrier (OFDM) system over sub-meter distances. A typical software-defined physical-layer transceiver system consists of frame generation, modulation, pulse shaping, pre-equalization, noise filtering, frame synchronization, post equalization, and demodulation. While pre-equalization accounts for the frequency-selective response of components, post-equalization mitigates ISI and the frequency-selective channel response. Several experimental testbeds that demonstrate multi-Gbps THz links over several distances have been reported, for both electronic (at 240 GHz [154], 300 GHz [155], 625 GHz [156] and 667 GHz [157]) and photonic systems (in the vicinity of 300 GHz [158]). Testbeds for above 1 THz operations, however, are still lacking.

There are several other notable THz experimental demonstrations. In [159], THz-band outdoor channel measurements are reported using double-directional frequency-domain channel sounding, where a large number of departure and arrival directions of significant energies are detected. In [160], Tbps speeds are demonstrated over THz LoS links; NLoS links through first-order reflections are shown to be feasible in cases of signal obstruction. This observation motivates the development of MIMO mechanisms that leverage spatial diversity in transmission, reception, and reflection. Furthermore, open-source large-scale distributed testbeds are being developed to facilitate experimental research in the mmWave and THz bands, such as MillimeTera [161]. In [162], a live streaming demonstration of an uncompressed 4K video using a photonics-based THz communication system (below 200 GHz) is reported, where error-free transmission is achieved at a distance of 1 m. Similarly, THz signals propagating through practical outdoor weather conditions and subject to indoor surface reflections are studied in [163]. In [164], it is also argued that with proper planning, THz communications can support outdoor communications.

V. THZ MODULATION SCHEMES AND WAVEFORM DESIGN

Designing efficient THz-specific waveforms and modulation schemes is a crucial step for unleashing the true powers of the THz band. On one hand, specific waveform designs can mitigate the limitations in THz sources and receivers. On the other hand, optimized and adaptive modulation schemes can make the best of the available spectrum. The choice of modulation schemes provides a compromise between low-complexity and high-rate physical layer configurations. In general, since carrier-based systems at higher frequencies tend to use larger spectrum bandwidths per channel, simple modulation schemes that require very low complexity digital demodulation are favored (binary phase shift keying and amplitude-shift keying). Nevertheless, novel modulation schemes and optimized multi-carrier waveform designs should be tailored for specific THz communication use cases.
A. Single-Carrier versus OFDM

As THz beams are expected to be narrow under high antenna gains, the corresponding delay spread is reduced (survival of a single path), and the channel should be flat. This may require a deviation from OFDM. In fact, OFDM is very complex to implement in the context of ultra-broadband and ultra-fast THz systems (complex transceivers with Tb/s digital processors still do not exist). The resultant strict frequency synchronization and high peak-to-average power ratio (PAPR) requirements also render OFDM ineffective in the THz Band. The limitations of DACs and ADCs, further, prevent the digital generation of multi-band orthogonal systems. SC modulation for above-90 GHz is motivated as a spectral- and energy-efficient solution for Tb/s wireless communications. Nevertheless, since there is an abundance of bandwidth in the THz band, non-overlapping, and perhaps equally spaced sub-windows, can prove to be efficient, as illustrated in several photonic THz experiments. Non-overlapping windows can be understood as SC modulation with some form of carrier aggregation, which is much less complex than OFDM, and would thus allow the use of high-frequency energy-efficient power amplifiers. Transmission can be conducted in parallel over these windows, where each carrier would occupy a small chunk of bandwidth that supports less data rate. This relaxes design requirements (simpler modulation and demodulation) and reduces energy consumption while retaining an overall high rate THz system. These benefits come at the cost of having to operate multiple modulators in parallel, where a very fast signal generator is required to switch between carriers. SC transmission is further advocated in [167]. In fact, SC has been proposed in WiFi 802.11ad (WiGig) for mmWave communications. The SC waveform can be complemented with simple continuous phase modulation (CPM) schemes such as continuous phase modulated SC frequency division multiple access (CPM SC-FDMA) and constrained envelope CPM (ceCPM-SC) [170]. SC can even provide larger power amplifier output power than cyclic prefix OFDM. Furthermore, SC transceivers are resilient to PN [125], [171], especially when combined with a PN-robust modulation scheme. In general, low-order modulations have low PAPR and are robust to PN. By tracking the phase of local oscillators at the transmitter and the receiver, using time-domain phase tracking reference signals, for example, SC can accurately estimate the PN at low complexity. Frequency selectivity can still emerge in THz systems, such as in indoor environments supporting a sufficient number of multipath components. The bandwidth of each THz transmission window (approximately 0.2 THz) can be much larger than the coherence bandwidth, which at 0.3 THz can be as low as 1 GHz in a multipath scenario and can reach 60 GHz with directional antennas. In fact, frequency selectivity increases with communication distance, pulse bandwidth, and center frequency. A frequency-selective system can also exist due to the behavior of the THz receivers themselves. Therefore, being a well-understood technology, OFDM can still be used for THz communications. In OFDM is proposed to enhance the spectral efficiency of 60 GHz systems. Furthermore, in hybrid THz MIMO-OFDM beamforming schemes with frequency-selective fading are proposed. In particular, analog beamforming is conducted via a normalized beamsteering codebook search algorithm, and digital beamforming is performed using a regularized channel inversion method. Alternative THz multi-carrier modulations other than OFDM are also being studied, such as wavelength division multiplexing (WDM) and Nyquist WDM [174].

In addition to the shortcomings of OFDM at THz frequencies, it is worth noting that special care should be given to mitigate the resultant Doppler effect. Under perfect frequency and time synchronization, the cyclic prefix length in OFDM is chosen so as to account for the delay spread in the system, and the OFDM symbol length is proportional to the inverse of the Doppler spread. Hence, at THz frequencies, the cyclic prefix is relatively larger for the same delay spread conditions. This problem is highlighted in [81] and is shown to complicate the OFDM design. Prospect solutions for this problem in the literature include beam-based Doppler frequency compensation schemes [175], [176].

B. Optimized Modulation Schemes

Modulation schemes can be further optimized to utilize the fragmented THz bandwidth, not only to mitigate the absorption effect but also to turn it into an advantage. Towards this end, and since the THz channel response is distance-dependent, distance-aware multi-carrier schemes that dynamically optimize transmission window allocations are proposed in [166]. Such schemes are shown to achieve Tb/s data rates, an order of magnitude higher than fixed-bandwidth modulation schemes, over medium-range communications (10 m). Nevertheless, such schemes come at the expense of a slightly increased complexity, as they typically require a control unit and a multi-carrier modulator. They also need a feedback path. Many more resources, however, can be dynamically optimized. For example, frequency allocation per AE is optimized in [61] to maximize capacity, as a function of the number of frequencies and AEs, as well as the antenna and array gains and beamsteering angles. Moreover, a pulse-based multi-wideband waveform design is optimized in [62] to enable communication over long-distance networks by adapting the power allocation criteria over a variable number of frames. The design incorporates pseudo-random time-hopping sequences and polarity randomization, and accounts for temporal broadening effects and delay spread; a communication range of 22.5 m and a 30 Gbps data rate are reported. A single-user and multi-user distance-aware bandwidth-adaptive resource allocation solution is further proposed in [177]; it supports a data rate of 100 Gbps over a 21 m distance. The unique relationship between distance and bandwidth are also exploited to enable a multi-carrier transmission [70].

A hierarchical modulation scheme is further proposed in [89] for a single-transmitter multiple-receiver system, that supports multiple data streams for various users at different distances by adapting the modulation order and symbol time. These optimization problems are extended to cases of large densities and user mobility. In such scenarios, multi-user interference is unavoidable. In particular, in [143], the
opportunistic use of resources under mobility is maximized to satisfy the constraints on humidity, distance, frequency bands, beamwidths, and antenna placement. Furthermore, a stochastic model of multi-user interference is proposed in [178], alongside modulation schemes that minimize the probability of collisions. THz OFDM adaptive distance- and bandwidth-dependent modulations are also considered in [93].

C. Pulse-Based Modulation

While continuous carrier-based transmission can be supported in the sub-THz range when the constraint on size is relaxed [36], [179], at true THz frequencies, carrier-based transmission is still challenging. For instance, it is not easy to generate more than short high-frequency pulses of few milli-watts with graphene at room temperature. Nevertheless, with large bandwidths, a reduction in spectral efficiency is acceptable, and pulse-based modulations can be used. Pulse-based SC on-off keying modulation spread in time (TS-OOK) that exchanges hundreds of femtosecond-long pulses between nano-devices is proposed in [180]. By assuming time-slotted operations with a time slot \( T \), the Gaussian pulse is expressed as

\[
p(t) = a \exp \left(-\frac{(t-b)^2}{2T_p^2}\right),
\]

where \( a, b, \) and \( T_p \) are the amplitude, center, and spread of the pulse, respectively (Gaussian pulse has a duration \( T_p < T \)). Conversely, the raised cosine pulse of the carrier-based system is expressed as

\[
q(t) = \sin c \left(t/T\right) \cos \left(\pi \alpha (t/T)\right) / \left(1 - 2\alpha^2 (t/T)^2\right),
\]

where \( \alpha \) is the roll-off factor \( (0 \leq \alpha < 1) \).

Pulse-based THz communications can achieve a Tbps data rate in nano-network scenarios [180]. They have also been used in ultra-wide-band impulse-radio systems [181] and free-space optics [182]. However, pulse-based systems are power-limited, especially in the case of extremely wideband signals. Wideband pulses are thus typically used to realize low-power, compact, and low-complexity sub-band transmissions. Nevertheless, by jointly optimizing the modulation and power allocation in an iterative manner, Tbps rates are demonstrated in indoor communication paradigms in [183], assuming realistic transmit, receive, and equalization filters, and under practical error rate constraints. An SC pulse-based approach is first addressed by optimizing the choice of modulation scheme, assuming a very long dispersive channel impulse response that accounts for ISI. Then, frequency division over multiple orthogonal sub-bands is considered, alongside efficient power allocation, to minimize the loss in the rate that is caused by finite-alphabet modulations. Furthermore, in [184], a parallel sequence spread spectrum (PSSS) is proposed as an alternative to OFDM for THz communications; it is shown to achieve 100 Gbps with simple receiver architectures that can almost be completely implemented in analog hardware.

VI. RECONFIGURABLE UM-MIMO ARRAYS

Given the quasi-optical behavior under LoS dominance, a THz MIMO channel is sparse with multi-user beamforming and low-rank with spatial multiplexing. As mentioned in Sec. [11] due to high directivity, and since beamforming is typically configured at the level of AEs within a SA, each SA is effectively detached from its neighboring SAs in a multi-user setting, and the role of baseband precoding reduces to defining the utilization of SAs, or to simply turning SAs on and off. In a point-to-point setup, however, the SA paths are highly correlated, and the channel is ill-conditioned. Nevertheless, good multiplexing gains are still achievable using sparse antenna arrays [185] that reduce spatial correlations in point-to-point LoS scenarios. The capacity of LoS MIMO uniform linear array channels is studied in [186] over all possible antenna arrangements. To enhance THz LoS channel conditions, spatial tuning techniques that optimize the separations between AEs can be applied [64].

A. THz Spatial Tuning

Antenna gains and array gains are pivotal for overcoming the path loss, where specific combinations of these gains are recommended for a specific communication distance and frequency of operation. While a mmWave system typically requires a footprint of few square centimeters for a few tens of antennas (not even sufficient to overcome the path loss over few tens of meters), a very large number of AEs can be embedded in few square millimeters at THz frequencies. Furthermore, depending on the communication range \( D \) and the SA separation \( \Delta \), three modes of operation can be distinguished: a mode where \( \Delta \) is large enough so that the channel paths are independent and the channel is always well-conditioned, a mode where \( D \) is large compared to \( \Delta \) and therefore the channel is ill-conditioned, and a mode where \( D \) is much larger than \( \Delta \), larger than the “Rayleigh” distance, where the channel is highly correlated. As long as the communication range is less than the Rayleigh distance in the second mode, \( \Delta \) can be adapted to enhance the channel conditions and achieve near-orthogonality. By finely tuning \( \Delta \), multiple data streams over eigenchannels can be transmitted. Both a shorter \( \lambda \) and a smaller \( D \) result in a smaller optimal separation of antennas \( \Delta_{opt} \), where in a perfectly aligned symmetric system \( M_t = N_t = M_r = N_r = M \) we have [64]

\[
\Delta_{opt} = \sqrt{\frac{\pi D \lambda}{M}}.
\]

for odd values of \( \pi \). Such optimizations can not be achieved in the third mode due to the limitation in physical array sizes. Alternative optimization schemes for antenna separations in LoS communications beyond the Rayleigh distance have been studied in [187], [188].

This adaptability in design, when combined with numerical optimization, is what we call “spatial tuning”. Spatial tuning is typically illustrated in the context of plasmonic antennas when having a sufficiently large uniform sheet of AEs. For example, a graphene-based sheet can consist of hundreds of uniformly-spaced active graphene elements mounted on a dielectric layer, which itself is mounted on a common metallic ground [55]. Hence, AEs can be contiguously placed over a 3D structure, and SAs can be virtually formed and adapted. For the desired communication range, a required number of AEs per SA is allocated. Then, the number of possible SA allocations, which is bounded by array dimensions and the number of RF chains, dictates the diversity/multiplexing gain.
Spatial tuning can be extended to include multicarrier design constraints. For instance, nano-antenna spacings in plasmonic antenna arrays can be reduced to \( \lambda_{\text{SPP}} \) while still avoiding the effects of mutual coupling. Mutual coupling in graphene-based THz antenna arrays is studied in [68], [189]. By using the couple mode theory, the impact of mutual-coupling on the response of nano-antennas is modeled via a coupling coefficient. Even at separations much less than \( \lambda \), near field mutual coupling is shown to be negligible. This is a promising realization that allows the practical implementation of compact THz systems in small footprints. In [189], the use of a frequency-selective surface structure that can be mounted between the array elements of an UM-MIMO array (behaving as a spatial filter) is proposed to reduce the mutual coupling effects further to negligible values.

Having AEs very close to each other, however, is not always beneficial, as this would reduce the spatial resolution and the achievable multiplexing gains. Furthermore, the inter-AE separation distance should not exceed \( \lambda/2 \), beyond which grating-lobe effects arise. By setting the separation between two active AEs to be \( \lambda/2 \), all the AEs in between would be idle. These antennas can be used for several purposes. Other than using them to increase the array gain, they can be configured to operate at different frequencies in a multicarrier scheme that supports more users with the same array footprint, as illustrated in Fig. 2. Alternatively, neighboring AEs can be configured to operate at the same frequency in a spatial oversampling setup [190], lowering the spatio-temporal frequency-domain region of support of plane waves [191]. The latter approach can be exploited for noise shaping, which results in a reduced noise figure and increased linearity.

### B. Index Modulation and Blind Parameter Estimation

Spatial modulation (SM) schemes for THz communications are promoted in [64], [192] as power- and spectrum-efficient solutions. The efficiency of SM at higher frequencies is highly dependent on the array design and channel conditions, as illustrated in [193]–[195] for mmWave systems. By mapping information bits to antenna locations in an adaptive manner, hierarchical SM solutions can be designed at the level of SAs or AEs. With SM, the number of bits that can be accommodated in a single channel use is

\[
N_b = \frac{\log_2(Mq)}{F} + \frac{\log_2(Q^2)}{F} + \log_2(|X|) \, .
\]

The transmitted binary vector over one symbol duration can be expressed as \( \mathbf{b} = [b_m b_q b_s] \in \{0,1\}^{N_b} \), where \( b_m \in \{0,1\}^{\log_2(Mq)} \) represent SA selection, \( b_q \in \{0,1\}^{\log_2(Q^2)} \) represent AE selection, and \( b_s \in \{0,1\}^{\log_2(|X|)} \) correspond to the actual QAM symbol. In such a design, the number of AEs and SAs, as well as the constellation size can be tuned for a desired bit rate. By enabling the selection of various combinations of antennas at the same time, a generalized index modulation scheme can be defined [196], [197]. Typical massive MIMO SM and generalized SM solutions [198], [199] should be revisited in the ultra-massive THz context. Furthermore, when enabling adaptive antenna-frequency maps, generic index modulation (IM) solutions that take full advantage of the available resources can be configured, in which information bits are also mapped to frequency allocations. The number of bits per channel use with IM can increase to

\[
N_b^{(\text{IM})} = \log_2 \left( \frac{\bar{F}}{F} \right) + \log_2 \left( \frac{S}{M^2 Q^2} \right) + \log_2 (|X|) \, ,
\]

where \( F \) is the total number of narrow frequency bands that are available, and \( \bar{F} \) and \( S \) are the number of frequencies that can be supported and the number of antennas that can be activated at a specific time, respectively. Even more generalized IM schemes can be realized by jointly designing the spatial and frequency bit maps. Such designs could prove to be particularly efficient in the THz band because a huge number of AEs can be fit in small footprints, and because the fragmented nature of the THz spectrum allows allocating multiple absorption-free spectral windows concurrently. However, the efficiency of these adaptive schemes is limited by the speed in which frequency hops can be executed. Nevertheless, changing the frequency of operation in the THz band can be achieved in...
a very fast manner without the need for changing the physical dimensions of the transmitting antennas. In particular, simple material doping or electrostatic bias are shown to change the Fermi energy of graphene, which dictates the frequency of operation. Software-defined plasmonic metamaterials are also a candidate solution, mainly for frequencies below 1 THz. For SC systems with IM, constant-envelope modulations such as CPM are shown to be power-efficient [200].

Such design compactness and flexibility can be further enhanced when complemented by THz-specific signal processing techniques at the receiver side. Instead of communicating transmission parameters with the receiver, blind parameter estimation can be conducted. For instance, in [201], a tertiary hypothesis test based on power comparison for antenna index and modulation mode detection is proposed and analyzed, alongside low-complexity frequency index detectors and modulation type estimators. Note that information bits can be assigned for the choice of modulation type as well. Despite the fact that modulation classification [202], [203] is a classical signal processing problem, its applicability to the THz-band [204] can prove to be particularly useful. Given the enormous possibilities of map-bit combinations, compressed sensing and machine learning techniques can be applied for detection and estimation purposes.

VII. BEAMFORMING AND PRECODING

As previously mentioned, beamforming and precoding are critical to overcome the high path losses at high frequencies and exploit the distance- and frequency-dependent characteristics of the THz channel. Since due to collimation maintaining alignment is much harder at THz frequencies, THz beamforming schemes should be fast. It can be argued that by the time 6G arrives, fully digital arrays operating at mmWave frequencies and below will be readily available and capable of achieving near-optimal beamforming performance. However, it is unlikely that this would be the case at sub-THz and THz frequencies. In line with state of the art, the motivation for hybrid beamforming in the THz realm is very similar to that in the mmWave realm [205]. There is a bottleneck in realizing prohibitively-complex and high-power-consuming fully-digital arrays. THz hybrid beamforming is motivated in [65], for example, where a distinction is made between the fully-connected configuration, in which one RF chain drives the entire antenna array, and the configuration in which an RF chain drives a disjoint subset of antennas, with a phase shifter per antenna (Fig. 3). At the receiver side, due to limited hardware and processing capabilities, a single RF chain is typically assumed to drive an antenna array.

A. THz Hybrid Beamforming

Given the limited power constraints of THz sources, the fully-connected configuration is expected to be power-aggressive, where the corresponding number of power-consuming combiners and phase shifters is very high [206]. Nevertheless, efficient fully-connected THz-band hybrid beamforming schemes can still be realized [207], [208]. The most popular THz beamforming and precoding designs follow the AoSA configuration of Sec. II. In such architectures, analog beamforming is configured using a large number of AEs per SA to achieve spatial energy focusing. A beam-steering codebook design can be used per RF chain since THz phase shifters can be digitally controlled [209]. This can be implemented through beam scanning, which ensures that the received signal power is largest for a specific user. Then, digital precoding at the level of SAs can be used to combat multi-user interference or to simply define the utilization of SAs when interference is negligible due to high directivity. The precoding problem is typically formulated as an optimization problem that minimizes the mean square error between the received signal and the transmitted symbols under the power constraint. Simple zero-forcing precoding at the baseband should be sufficient in several THz scenarios. In highly correlated point-to-point THz links, however, better-performing efficient nonlinear precoding techniques are required, such as block multi-diagonalization [210], [211]. The energy efficiency of the AoSA configuration is better than that of the fully-connected one [212]. This is further emphasized when taking into consideration the nonlinear system power consumption model and insertion losses.

The problem of finding the best allocation of SAs to enhance spectral efficiency, i.e., hybrid precoding with dynamic antenna grouping, is an important problem in THz UM-MIMO beamforming. Towards this end, switches can be inserted in a dynamic AoSA (DAoSA) architecture to tune the connections between SAs and RF chains [213], [216]. This flexibility is optimally realized in a fully-connected architecture, in which establishing fully-dynamic connections requires an exhaustive search over all possible connections between RF chains and SAs. Moreover, the complexity and power consumption of such fully-connected systems is prohibitive at high frequencies and large dimensions (thousands of switches). Towards exploiting this trade-off between spectral efficiency and power consumption, near-optimal and low-complexity THz hybrid precoding algorithms are proposed in [110]. The design problem is divided into two sub-problems: The hybrid DAoSA precoding problem, and the switch selection problem. The system model corresponding to hybrid precoding in DAoSA architectures is a modification of (1), where the precoding and combining matrices $W_f^H$ and $W_d^H$ are decomposed into their digital and analog components as

$$y = \sqrt{p}C_D^H C_A^H H P_A P_D x + C_D^H C_A^H n,$$

where $p$ is the transmitting power and $y$ and $x$ here have the dimension $N_s \times 1$, with $N_s$ being the actual number of data streams, $H \in C_{M_r, N_s, Q^2 \times M_t, N_r}^H$, and $n \in C_{M_t, N_r, Q^2 \times N_r}$. Furthermore, $C_A \in C_{M_r, N_s, Q^2 \times M_t, N_r}$ and $C_D \in C_{M_r, N_t, N_r}$ are the analog and digital combining matrices, and $P_A \in C_{M_t, N_t, Q^2 \times M_t, N_t}$ and $P_D \in C_{M_t, N_t, N_t}$ are the analog and digital precoding matrices, respectively. The achievable rate of this system model (subject to optimization) can be expressed as

$$R = \log_2 \left( \left| I_{M_t, N_r, Q^2} + \frac{p}{N_t g^2} HP_D (P_A P_D)^H H^H \right| \right),$$

where $I_{M_t, N_r, Q^2}$ is the identity matrix of size $M_t N_r Q^2$ and $|\cdot|$
Fig. 3: High-frequency hybrid beamforming architectures: Fully-connected, fixed AoSAs, and dynamic switching.

denotes the matrix determinant operator.

Hybrid beamforming often entails user grouping and SA selection. In [112], a THz multi-carrier distance-dependent hybrid beamforming scheme is proposed, in which user grouping is achieved through analog beamforming, alongside digital beamforming, power allocation, and SA selection mechanisms. The proposed solution allows users of different user groups to share frequencies while avoiding interference in the analog domain. Users in the same group, however, are assigned orthogonal frequencies based on a distance-aware multi-carrier scheme. SAs then get assigned to the data streams of a user group in the digital domain. The same authors address the hybrid beamforming problem for THz indoor scenarios in [114]. A bound on the ergodic capacity is derived, and the relation between the required size and number of SAs and the communication distance is established. In general, the spectral efficiency gap between hybrid and digital precoding is smaller when the channel is sparser [217], which makes hybrid schemes suitable for multipath-limited THz channels.

Developing novel wideband hybrid beamforming schemes for THz communications is also very important. For instance, in [218], the authors propose an OFDM-based normalized codebook search algorithm for beamsteering and beamforming in the analog domain and a regularized channel inversion method for precoding in the digital domain. Two digital beamformers are used in a three-stage scheme to account for the loss in performance due to hardware constraints and the difference between subcarriers. Similarly, in [81], a beam division multiple access scheme is proposed for wideband massive MIMO sub-THz systems, which schedules a mutually non-overlapping subset of beams for each user. The algorithm is based on per-beam synchronization in time and frequency, taking into account the delay and Doppler frequency spreads, the latter of which are orders-of-magnitude larger at THz frequencies. Note that an inaccurate narrowband assumption, in which the precoding and combining matrices of a wideband system are designed for a specific carrier frequency, induces the effect of beam split in THz communications. This phenomenon is mitigated in [219] by using a THz-specific delay-phase controlled precoding mechanism, in which time-delay components are introduced between the RF chains and the phase shifters. The addition of such components has the potential to create frequency-dependent beams that are uniformly aligned with the spatial directions over the entire bandwidth.

In addition to the aforementioned considerations for beamforming in the THz band, all of which are extensions to similar considerations in the mmWave band, recent studies are considering novel THz-specific beamforming schemes based on novel THz circuitry. For instance, in [220], a graphene-based dense antenna array architecture is proposed, in which each element is integrated by a THz plasmonic source, direct signal modulator, and nano-antenna. For such an architecture, novel dynamic beamforming schemes at the level of single elements, as well as the level of the integrated array, are proposed, where full phase and amplitude weight control can be achieved by simply tuning the Fermi energy of the modulator and AE. The authors propose a codebook design for Fermi energy tuning that results in reasonably accurate beamforming and beamsteering; it is shown that the power-density of the array increases non-linearly with its size. Note that controllable THz frequency-dependent phase shifters can also be realized via low-loss integrally gated transmission lines [209], the length of which determines the signal travel time, and hence the phase shift. Furthermore, graphene/liquid crystals have been proposed for magnet- or voltage-controlled THz phase shifters [221], [222]. Since they are digitally controlled, such phase shifters only generate quantized angles.

Several other beamforming considerations need to be considered for future THz networks. For example, in a cell-free massive MIMO scenario [223], distributed access points can each provide an excess of 100 GHz to a user, especially under low mobility. Dense deployment of access points can guarantee short THz communication distances, even under significant blockage. Mobility and blockage are addressed in [224] in the context of network massive MIMO scenarios in the mmWave and THz bands, where per-beam synchronization is proposed to mitigate the channel Doppler and delay dispersion, and precoding beam domain power allocation is reduced to a network sum-rate maximization problem.

B. One-Bit Precoding

At higher frequencies and ultra-massive dimensions, the circuit power consumption, hardware complexity, and system cost significantly increase. The dominant sources of power consumption are analog-to-digital converters (ADCs) in the uplink and digital-to-analog converters (DACs) in the down-
link. The power dissipation in converters scales exponentially in the number of resolution bits, and state of the art DACs and ADCs can only achieve 100 Gigasamples-per-second rates [223]. Furthermore, the capacity requirements on the fronthaul interconnect links are also severe in large MIMO systems. Jointly reducing system costs, power consumption, and interconnect bandwidth with minimal performance degradation remains a challenge. As an alternative to reducing the number of converters using hybrid beamforming, the bit resolutions can be reduced through coarse quantization. The latter approach has the extra benefit of lowering the linearity and noise requirements, which is crucial in THz settings. In the extreme case of one-bit quantization [226], only simple comparators are required, and automatic gain control circuits are no longer required. Noting that for high amplitude resolutions, the power consumption of ADCs grows quadratically with the sampling rate, a one-bit quantization solution is proposed in [227] for sub-THz wideband systems, where the amplitude resolution is reduced while accounting for that by temporal oversampling.

By modifying our system model (1), the precoded transmitted symbol vector can be expressed as \( \bar{x} = [\bar{x}_1, \ldots, \bar{x}_b, \ldots, \bar{x}_{M,N}]^T \in \mathbb{X}^{M,N} \times 1 \), where under finite-precision, the \( b \)th symbol of \( \bar{x} \), \( \bar{x}_b = I_R + jI_I \in \mathbb{X} \), has quantized in-phase and quadrature components, i.e., \( I_R, I_I \in \mathcal{L} \), where \( \mathcal{L} = \{l_0, l_1, \ldots, l_{L-1}\} \) is the set of possible quantization labels and \( \mathbb{X} = \mathcal{L} \times \mathcal{L} \). For 1-bit quantization, we have \( \mathcal{L} = |\mathcal{L}| = 2 \). Prior to precoding, the symbol vector \( x \) is obtained by mapping the information bits to the original constellation \( \mathbb{X} \). The base station then uses the knowledge of \( H \) to precode \( x \) into \( \bar{x} \). Note that \( x \) and \( \bar{x} \) need not be of the same size. With coarse quantization, there is an additional distortion factor due to finite precoder outputs. Since optimal precoding is exhaustive due to the cardinality of \( \mathbb{X}^{M,N} \), in THz UMiMIMO systems, only linear quantized precoders [228] are feasible, or perhaps very few optimized low-complexity non-linear quantized precoders as well [229]. Analyzing the system performance under quantization is typically conducted using the Bussgang decomposition [230].

The performance of THz indoor one-bit distance-aware multi-carrier systems is investigated in [231] for a hybrid precoding AoSA architecture. The achievable rate is shown to be insensitive to changes in transmit power, and single-user transmission is shown to be robust to the phase uncertainties in large antenna arrays. The best beamsteering phase shifter direction, further, is shown to be that of the LoS path.

C. THz NOMA

Non-orthogonal multiple access (NOMA) techniques [232], [233] have been recently proposed to combat the loss of spectral efficiency in orthogonal multiple access schemes, especially when the resources are allocated to users with poor channel conditions. The lack in spectral efficiency, however, is not a primary bottleneck for THz communications, given the largely available bandwidths and the fact that higher spatial resolution in beamforming limits the need for multiple access schemes. Yet, any additional spectral efficiency enhancement technique is welcome if its additional complexity cost is limited. NOMA at higher frequencies [234] is thus more likely to be conducted over point-to-point doubly-massive MIMO links. In such scenarios, the concept of multiple access reduces to superposition coding of multiple data streams over a single link. Nevertheless, calling the resultant configuration NOMA is not a misnomer, because each THz beam can still be configured to serve multiple users. In the latter scenario, the role of NOMA can prove to be essential in mitigating the hardware constraints in THz devices, which limit the beamforming capabilities.

For power-domain NOMA, and following our system model, multiple data streams can be concurrently sent via superposition coding over different combinations of transmitting and receiving SAs that form overlapping effective channel matrices. Assume, for ease of construction, that the superposition of data symbols occurs at the lower layers of the MIMO channel matrix. Denote by \( S \) the set of power-domain multiplexed data streams of dimensions \( S_i, i = 1, \ldots, |S| \), such that \( S_i \geq S_{i+1} \) and \( S_1 = M_t N_t \). The multiplexed transmitted symbol vector \( x_i = [x_n^i, x_n^i, \ldots, x_n^i] \in \mathcal{X}^{S_i} \times 1 \) is allocated to the contiguous set of antennas \( N - S_i + 1 \) to \( M_t N_t \). We thus have the effective channel matrices \( H_i \in \mathbb{C}^{M_t, N_t \times S_i} \) to be comprised of the columns \( M_t N_t - S_i + 1, M_t N_t - S_i + 2, \ldots, M_t N_t \) of \( H \). The equivalent baseband input-output system relation can then be expressed as

\[
y = \sum_{i=1}^{|S|} H_i x_i + n_i
\]

in which NOMA is achieved by assigning different power levels to the multiplexed transmitted symbol vectors. For example, we can allocate a higher power level \( p_i \) to the symbol vectors \( i \) of smaller dimension, i.e., \( p_1 < p_{i+1} \). Each symbol \( x_n_i \) thus belongs to a scaled complex constellation \( \mathcal{X}^i \) (\( E[x_n_i^2] = p_i \)), and we have \( x_i \in \mathcal{X}^i \), the lattice that includes all possible symbol vectors generated by \( S_i \mathcal{X}^i \) constellations. Note that with higher degrees of reconfigurability in THz antenna arrays, generalized coding schemes encompassing antenna selection, frequency, and power allocation can be easily realized. However, low-complexity detection and decoding schemes should complement such designs, in particular, efficient successive interference cancellation at the receiver.

When sufficient multipath components exist, perhaps in scenarios where lower antenna gains are required (indoor sub-THz scenarios, for example), conventional single-cell multi-user MIMO-NOMA settings can still be realized. Assume that the cellular users are divided into two groups, where the first group of users are uniformly distributed in an inner disk (\( C_1 \)) centered at the base station and of radius \( R_N \), while the second group of users are uniformly distributed in an outer disk (\( C_2 \)) from \( R_N \) to \( R_C \). A base station with \( N \) transmitting antennas simultaneously services two users, user 1 with \( M_1 \) antennas in the inner disk (\( C_1 \)) and user 2 with \( M_2 \) antennas in the outer disk (\( C_2 \)), in the same frequency and time slot via power-domain superposition coding. The received vectors \( y_1 \) and \( y_2 \) at user 1 and user 2 can be expressed as

\[
y_1 = H_1 x_1 + H_1 x_2 + n_1
\]

\[
y_2 = H_2 x_1 + H_2 x_2 + n_2
\]
where $\mathbf{x}_1$ and $\mathbf{x}_2$ are the power-multiplexed transmitted symbol vectors and $\mathbf{H}_1$ and $\mathbf{H}_2$ are the corresponding channel matrices. In this scenario, NOMA is achieved by clustering users from the inner disk $C_1$ with users from the outer disk $C_2$ and assigning different power levels to the multiplexed transmitted symbol vectors. Energy efficiency in a THz MIMO-NOMA system is addressed in [235] by optimizing the user clustering, hybrid precoding, and power allocation mechanisms.

VIII. THZ Baseband Signal Processing

Efficient baseband signal processing is key for mitigating the impairments of novel THz-band devices and enabling operations beyond 100 Gbps [236]. The true bottleneck at the baseband is the lack of energy-efficient transceivers that can approach a Tb/s data rate [59], where the sampling frequency is still in the order of 100 gigasamples/second in state of the art ADCs and DACs. Efficient signal processing across all baseband blocks is required to fill this gap. In fact, due to the diminishing effect of Moore’s Law, silicon scaling will provide limited improvements in baseband computations and chip power density. Hence, incorporating application-specific integrated circuit (ASIC) architectures and baseband algorithms in a holistic framework is the only option for realizing Tb/s operations in mobile terminals. Therefore, the enhancement in the energy efficiency of new THz devices should be complemented with joint algorithm and architecture co-optimization. In the following sub-sections, we highlight recent advancements in low-complexity channel estimation, channel coding, and data detection schemes that optimize the overall power consumption, area efficiency, and latency.

A. Channel Estimation

Channel estimation in the THz band is very challenging in mobile scenarios, where accurate channel state information (CSI) is required for beamforming mechanisms and for accurately directing beams to avoid misalignment issues. Accurate CSI is particularly important in the absence of an LoS path. Furthermore, frequent channel estimation might also be required for fixed LoS point-to-point THz links; this is because, at the micrometer wavelength scale, slight variations in the environment can introduce significant channel estimation errors. Classical channel estimation techniques should also be revisited by taking into account low-resolution quantization and hybrid analog and digital designs. Several techniques can be considered to reduce the complexity of THz-band channel estimation, such as fast channel tracking algorithms, lower-frequency channel approximations (exploiting outband signals), compressive-sensing-based techniques, and learning-based techniques, to name a few.

Using compressive-sensing techniques for sparse channel recovery in THz channel estimation is inspired by the successful use of these techniques in mmWave communications [237], where channels are even less sparse. For instance, in [238], approximate message passing (based on belief propagation in graphical models) and iterative hard-thresholding are argued to be an efficient compressed-sensing-based technique for THz channel estimation. Learning-based THz channel estimation schemes are most efficient at higher dimensionalities. Deep kernel learning based on the Gaussian process regression is explored in [239] for multi-user channel estimation in UM-MIMO systems over 0.06–10 THz and is shown to be efficient.

Despite channel sparsity, the real-time THz channel estimation complexity overhead can be significant in a dense multi-user wideband scenario with many paths; a large number of measurements might be required for compressive-sensing-based estimation. Towards this end, traditional minimum mean square error (MMSE) and least square channel estimation methods can be used to estimate the second-order statistics of THz channels [65]. Furthermore, joint activity detection and channel estimation is an efficient technique to reduce the use of pilots and the complexity of computations in wideband random massive-access THz systems [240]. Fast channel tracking, further, is an alternative approach to reduce the channel estimation overhead in high-mobility scenarios, as illustrated in [241] for THz beamspace massive MIMO.

B. Channel Coding

Towards bridging the Tb/s gap in baseband signal processing, data detection and decoding algorithms and architectures need to be of low latency and of high energy-efficiency and throughput. They should also be highly-parallelizable (spatial and functional parallelism) and should possess large data locality and structural regularity. Channel coding, in particular, is hitting the implementation wall, as it is the most computationally demanding baseband process [60]. The three main candidate coding schemes for 6G are Turbo, low-density parity-check (LDPC), and Polar codes. While Turbo and LDPC decoders are both executed on data-flow graphs, Turbo decoding is inherently serial, and LDPC decoding is inherently parallel. Polar decoding, on the other hand, is typically performed on a tree structure and is inherently serial. Due to their parallel nature, LDPC decoders provide higher throughput [242]. However, Polar and Turbo codes provide better flexibility in code rates and block sizes [59], which is much required in 6G.

A modular framework for generating and evaluating high-throughput Polar code decoders is presented in [243], where soft cancellation algorithms are explored. Achieving a Tb/s throughput with Polar codes is also addressed in [244], where low-latency majority logic and low-complexity successive cancellation are combined for decoding, alongside an adaptive quantization scheme for log-likelihood-ratios (LLRs). It is demonstrated that this scheme achieves Tb/s in a 7 nm technology implementation while occupying a 10 mm$^2$ chip area and consuming a 0.37 W power. Turbo codes have also come a long way towards beyond 100 Gbps operations [245, 247]. While such advances in coding schemes serve the ultimate goal of THz communications, which is achieving Tb/s operations, they are blind to the inherent characteristics of THz channels. Nevertheless, the THz channel can be taken into consideration in MIMO detection schemes, as well as in joint modulation, coding, and detection algorithms and architectures.
C. Data Detection

Although channel code decoding is the most computationally demanding baseband processing block, data detection also adds a significant computational burden, especially in doubly-massive MIMO systems. In conventional massive MIMO systems at lower frequencies, with a very large number of antennas at the base station and a few antennas at the receiving equipment, channel hardening occurs. With channel hardening, simple linear detection schemes such as zero-forcing and MMSE can achieve near-optimal performance. This is not the case in THz systems, where symmetric doubly-massive MIMO systems are common [67], especially because compact large THz antenna arrays can be embedded in the user equipment. In the latter scenario, the channel tends to be highly correlated, especially under THz LoS-dominance. Inter-channel interference prohibits using simple linear detection schemes that fail to decouple spatial streams and result in noise amplification.

More sophisticated non-linear detection schemes should thus be considered. However, the complexity of optimal non-linear detection schemes that achieve near-maximum likelihood performance is prohibitive at large dimensions. Therefore, novel THz-specific MIMO detectors which can achieve near-optimal performance with reasonable complexity are required. Conventional near-optimal detectors mainly replace the full-lattice search over all candidate transmit vectors in maximum likelihood detection with a reduced search over a reduced space of vectors that are closer to the truly transmitted vector (from a Hamming distance perspective). Such reduced-complexity detectors are mainly variations of sphere decoding schemes [248], [249]. Despite the fact that sequential processing in sphere decoding results in variable complexity and limits parallelism, several algorithmic and architectural optimizations have been proposed [250] to fix its complexity [251]. In [252], [253], the complexity of sphere decoding is reduced by casting memory-bound computations into compute-bound operations, and real-time processing is maintained by using graphics processing units.

However, even fixed-complexity sphere decoding is prohibitively complex if used for UM-MIMO detection. Recently, several detection algorithms that are suitable for large doubly-massive MIMO systems have been proposed. Such algorithms are mainly based on local search criteria [254], heuristic tabu search algorithms [255], message passing on graphical models [256], Monte Carlo sampling [257], and lattice reduction [258]. The Bell laboratories layered space-time (BLAST) detection algorithm is also modified to support ultra-high data rates in massive MIMO scenarios in [259]. Furthermore, perturbation-based regularizations can be used for equalization with ill-conditioned channels [260].

One family of detectors, in particular, that can achieve a good trade-off between performance and complexity in large highly-correlated MIMO channels, is the family of subspace detectors [66], [261]–[264]. These detectors mainly exploit channel puncturing to reduce complexity and enhance parallelism. In particular, since the computational cost of MIMO detectors is proportional to the number of nonzero elements in a channel matrix (most detectors involve back-substitution and slicing operations), by puncturing the channel into a specific structure, the detection process can be simplified and accelerated. More importantly, subspace detectors can break the interconnection between spatial streams, which significantly enhances parallelism at a marginal cost of multiple channel decompositions. Channel puncturing can be generalized to channel shortening, which can be used to mitigate ISI in SC THz systems [265]. The performance gap between optimal channel shortening (from an information-theoretic perspective) and channel puncturing can be covered by adding MMSE prefilter and channel-gain compensation stages [266], [267].

Other THz-specific data detectors include envelope- or energy-based detectors, which enable direct baseband operations without frequency down-conversion, thus bypassing phase impairments. Interference cancellation algorithms, further, can mitigate the effect of CFO. Compressed detection with orthogonal matching pursuit is also considered for sparse pulse-based multipath THz communications [268]. For a broadband SC THz system, MMSE precoding and detection is explored in [269] by assuming sparse channel matrices. Finite alphabet equalization [270], which reduces the complexity, power consumption, and circuit area (by coarsely quantizing the equalization matrix), can also prove to be useful for THz scenarios. Indoor THz communications Tb/s rates under finite alphabets are demonstrated in [183], where a frequency-division scheme of multiple sub-bands is utilized to relax the requirements on ADCs and DACs.

D. Joint Coding, Modulation, and Detection

One way of incorporating the effect of the THz channel in decoding is to consider iterative detection and decoding schemes. In particular, if MIMO detectors generate soft-output LLRs, these LLRs can be fed as soft inputs to the decoder, the output LLRs of which can then be fed again as soft inputs to the detector [271]. However, the extra complexity in iterations and in computing soft-output values in the detector should be taken into consideration. In the particular case of MIMO detection with Polar code decoding, iterations can be configured per stream-detection, after each decoding successive cancellation; the output of every step in the decoder can thus be used to enhance channel equalization.

Parallelizable detectors are favored for such designs. However, with parallelizability, the transmission vectors per stream would typically consist of a smaller number of bits. While Polar and Turbo decoders can cope with that, LDPC decoders might not perform well. Larger modulation types can be considered to increase the number of bits per stream for better decoding, 1024 QAM and beyond [272], for example. However, THz systems do not perform well with higher-order modulations, due to the increased complexity and the effect of PN. Alternatively, deeply-pipelined MIMO architectures can be used to aggregate data for the decoders, but this comes at the expense of reduced throughput. Nevertheless, since different transmission vectors within a decoder block can be independent, multiple detectors can operate in parallel.

In addition to joint channel coding and data detection, joint modulation design and coding can make the best use of
resources, especially in multi-wideband THz communications. Efficient probabilistic signal shaping techniques can be used in this regard, especially in highly reconfigurable UM-MIMO systems. In [276], the authors demonstrate that due to the peculiarity of noise in the THz band, proper use of channel codes can increase single-user and network capacity beyond classical networks with AWGN. Furthermore, THz-specific coding schemes are also being introduced at the network level. For instance, systematic random linear network coding (sRLNC) is proposed in [277], [278] for generic THz systems, in which coded low rate channels carry redundant information from parallel high rate channels. By tuning the transmission and code rates, the number of channels, and the modulation format, fault-tolerant high-throughput THz communications can be supported at different communication ranges. Also, in the context of THz index modulation, joint data detection and parameter estimation can be executed at the receiver side [279]; in the particular case of subspace detection, the entire process can be parallelized [280], [281].

Finally, it is worth noting that THz digital baseband operations of any complexity might be prohibitive in some use cases. Therefore, designing an all-analog THz baseband chain is a reasonable solution. Such design should include, in addition to analog modulation and detection schemes, all-analog decoders [282], which are based on soft-output transmission and detection (bit LLRs represented by currents and voltages). All-analog MIMO schemes can be realized using a continuous mapping scheme, which is noise limited but might perform well under interference.

IX. RECONFIGURABLE SURFACES

One of the recent hot research topics in wireless communications is making the environment intelligent and programmable for communication purposes [283]. Towards this end, a straightforward approach is to install large active arrays of AEs, also known as active large intelligent surfaces (LISs) on indoor and outdoor walls and other structures. This approach is a special case of UM-MIMO, and it is suitable for THz scenarios. In particular, with little restrictions on how to spread antennas over a surface, the mutual coupling effects can be avoided, and channel correlation can be reduced in LoS environments (spatial tuning of Sec. can be easily configured). Furthermore, channel estimation and feedback mechanisms can be easily achieved in active LIS setups, which is important to achieving low-latency THz communications. Another form of active large surfaces is the concept of holographic MIMO surfaces [287]. Since holographic MIMO borrows techniques from the optical domain, its implementation in the THz band should be more convenient than at lower frequencies.

It is the concept of passive IRSs [288], however, that is gaining most of the attention. IRSs are typically implemented using reflective arrays or software-defined metasurfaces, which introduce phase shifts at the level of reflecting elements (non-specular reflections), to focus and scale-up the power of reflected signals, and to steer beams into a particular direction. All that can be achieved without requiring complex encoding and decoding schemes or additional RF operations [290]. Note that phase shifts can also be implicitly achieved by tuning the impedance or length of delay lines. Metasurface elements can be much smaller than those of reflect arrays (which typically follow the half-wavelength rule), and they hence support more functionalities, such as polarization manipulation and absorption of incident waves. Reflections from tiny reflecting elements (of sub-wavelength size) form scattering in all directions, the collaborative effect of which results in beamforming. The control complexity of metasurfaces, however, can be higher. Compared to LISs, both types of IRSs are passive (note that LIS and IRS are used interchangeably in the literature). Nevertheless, IRSs should be electronically active at some level, perhaps to send pilots and for operational purposes.

IRS systems are particularly favorable in the THz band, where they can introduce controlled scattering to extend the very limited achievable communication distances and to enable multicasting. Hence, IRSs can add synthetic multipath components to enhance the performance of multipath-limited THz systems. The necessity for IRSs operating at THz frequencies also arises from the fact that regular coherent large antenna arrays are not easily achieved with the very small size of AEs. Even the relaying technology at THz is not mature. IRS systems are thus a viable solution. Another argument for IRS THz deployments stems from the limitations of surfaces themselves. In particular, it is argued that for IRSs to achieve SNRs comparable to those of massive MIMO, or to beat a classical half-duplex relay system, a large number of reflecting elements is required. This could result in physically-large arrays that are harder to deploy and that are subject to beam-squinting [291]. At THz frequencies, however, an electronically-large IRS compared to the operating wavelength can be achieved in very small footprints, which means that dense THz-band IRS deployments for short-range communications can be easily achieved. In addition to regular IRS functionalities, at THz frequencies, reflect arrays can be used at the transmitter to generate and direct a THz beam that is excited by a close THz source [105].

In addition to the THz-specific communication system considerations, several material properties favor THz-band IRS deployments. Metamaterials and metasurfaces operating in the THz band can provide the flexibility for generating orbit angular momentum and polarization conversion [292]. Furthermore, graphene-based metasurfaces can control the chemical potential of reflecting elements via electrostatic biasing, which varies the complex conductivity to achieve phase control [293]. Graphene-based digital metasurfaces combining both reconfigurable and digital approaches are studied in [294], where beamsteering is achieved by dynamically adjusting a phase gradient along the metasurface plane. A graphene-based metasurface is also proposed in [295], where a two-dimensional periodic array of graphene meta-atoms is shown to guarantee a wideband perfect-absorption polarization-insensitive reconfigurable behaviour at THz frequencies. In [296], the proposed curvilinear THz metasurface design is argued to be independent of the geometry and the frequency. THz metasurface-based beamsteering techniques are thus expected...
to achieve wide-angle ranges at high compactness and a light weight. The concept of HyperSurfaces, further, is proposed in [307] for THz communications. HyperSurfaces are composed of a stack of virtual and physical components that can enforce lens effects and custom reflections per tile. In addition to graphene, thermally or electrically tunable vanadium dioxide and liquid crystals, as well as micro-electromechanical systems, have also been considered as candidates for efficient THz steering technologies [302]. Novel intelligent plasmonic antenna array designs for transmission, reception, reflection, and waveguiding of multipath THz signals are further studied in [298], where an end-to-end physical model is developed. Nevertheless, future metasurfaces should support higher reconfigurability and sensing accuracy to support spatially-sensitive THz communications. A novel distributed control process, perhaps aided by optical internetworking, can guarantee fast adaptation [297].

Performance analysis studies for IRS-assisted communications include [291], [299]–[305]. The corresponding performance limits at high frequencies [105], [306]–[309], however, are still lacking. Most analytical studies consider lower-frequency scenarios and assume a downlink multi-user system model, in which an IRS is in the LoS of a base station assisting in reaching multiple users, each having a small number of antennas. For such setups, in [299], the minimum achievable signal-to-interference-plus-noise ratio (SINR) is studied, both when the channel between the IRS and the base station is full-rank or rank-one. The optimality of passive beamforming in IRSs is studied in [301], where a novel modulation scheme that avoids interference with existing users is proposed, alongside a resource allocation algorithm. Opportunistic scheduling is further studied in [300] as a means to achieve good multi-user diversity gains in spatially correlated LoS scenarios. Furthermore, the effect of random blockages in large-scale surface deployments is studied in [310] (blockage mitigation is crucial at THz frequencies). However, the THz multi-user channel is very sparse, and most of the analytical frameworks have to be revised in the THz context.

Several attempts for channel modeling in IRS-assisted scenarios at high frequencies are noted. For instance, in [311], a 3D channel model for indoor hypersurface-assisted communications at 60 GHz is developed. Channel estimation for IRS-assisted THz communications is studied along with hybrid beamforming in [308], where cooperative channel estimation is achieved via beam training and by exploiting the advantages of high dimensionalities and poor scattering at THz frequencies. End-to-end 3D channel modeling of radiation patterns from graphene-based reflectarrays at true THz frequencies is also presented in [105]. Furthermore, in [306], indoor IRS-assisted THz communications are studied, where a near-optimal low-complexity phase shift search scheme is proposed as an alternative to the complex exhaustive search. Beamforming in THz scenarios incorporating both graphene-based UM-MIMO arrays and metasurfaces is also studied in [312], showcasing the potential of combining these two technologies. Note that IRS-based index modulation schemes [313] can also be very efficient at THz frequencies. THz RIS systems can further assist virtual reality applications, as illustrated in [314].

For an IRS consisting of \( M \times N \) reflecting elements, the IRS-assisted NLoS communications system model, assuming AoSAs being placed at the transmitter and the receiver, is a simple extension of (1) which can be expressed as

\[
y = W_{r}^{H} \left( H_{IR} \Phi H_{TI} \right) W_{t}^{H} x + \tilde{n},
\]

where \( H_{IR} \in \mathbb{C}^{M_{t} \times M_{r} \times N_{t} \times N_{r}} \) is the channel between the IRS and the receiving array, \( H_{TI} \in \mathbb{C}^{M_{r} \times M_{t} \times N_{t} \times N_{r}} \) is the channel between the transmitting array and the IRS, \( \tilde{n} \) is the equivalent noise vector at the receiver, and \( \Phi = \text{diag} \left( \beta_{11}, \ldots, \beta_{M_{t}N_{t}} \right) \in \mathbb{C}^{M_{t} \times M_{r} \times N_{t} \times N_{r}} \) is the diagonal matrix that comprises the gains (\( \beta \)'s) and phase shifts (\( \Phi \)'s) at each IRS element.

Unlike at lower frequencies, where channel hardening effects can arise with large IRSs [303], at high frequencies, the channel is highly correlated and of low-rank. Hence, in addition to increasing the signal strength, IRSs operating at high frequencies should enhance the system performance by increasing the overall channel rank and suppressing interference. The authors in [307] demonstrate how an IRS can be used to increase the channel rank, leading to substantial capacity gains. To illustrate this issue at THz frequencies, we extend the concept of spatial tuning in Sec. VI-A to IRS-assisted THz NLoS environments. In a simplified proof-of-concept binary IRS operation, we assume that each reflecting element can either fully absorb an incident signal or reflect it towards a target direction. Hence, by controlling which element to reflect, a spatial degree of freedom is added at the
IRS level, which could enhance the multiplexing gain of the NLoS system, but at the expense of a reduced total reflected power. Global solutions can be derived by jointly optimizing $\Delta_1$, $\Delta_2$, and $\Delta_3$, the inter-SA spacing at the transmitting array, inter-reflect-element spacing at the intermediate IRS, and the inter-SA spacing at the receiving array, respectively, as illustrated in Fig. 4. For this special case, the channel matrices (less the molecular absorption factor) can be approximated as

$$H_{\text{IR}} \approx \frac{c}{4\pi f D_2} e^{-j2\pi D_2 \frac{\Delta}{c}} \left( \begin{array}{cccc} \omega_{1,1} & \omega_{1,2} & \cdots & \omega_{1,M,V_i} \\ \omega_{2,1} & \omega_{2,2} & \cdots & \omega_{2,M,V_i} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{N_r,V_i,1} & \omega_{N_r,V_i,2} & \cdots & \omega_{N_r,V_i,M,V_i} \end{array} \right)$$

$$H_{\text{TI}} \approx \frac{c}{4\pi f D_1} e^{-j2\pi D_1 \frac{\Delta}{c}} \left( \begin{array}{cccc} \tilde{\omega}_{1,1} & \tilde{\omega}_{1,2} & \cdots & \tilde{\omega}_{1,N_r,V_i} \\ \tilde{\omega}_{2,1} & \tilde{\omega}_{2,2} & \cdots & \tilde{\omega}_{2,N_r,V_i} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\omega}_{M,V_i,1} & \tilde{\omega}_{M,V_i,2} & \cdots & \tilde{\omega}_{M,V_i,N_r,V_i} \end{array} \right)$$

where $\omega_{n_r,V_i,m,V_i} = e^{-j\Psi_1(\Delta_n,\Delta_1)/2D_2}$ and $\tilde{\omega}_{m,V_i,n,V_i} = e^{-j\Psi_2(\Delta_n,\Delta_1)/2D_1}$, with $\Psi_1$ and $\Psi_2$ being functions of the specific geometry and coordinate system, and $D_1$ and $D_2$ being the distances between the centers of the IRS and the transmitting array and the IRS and the receiving array, respectively.

Compressive sensing and machine learning techniques for IRS-assisted THz communications are promising. For instance, in [315], the training overhead for channel estimation and the baseband hardware complexity are both reduced by assuming a sparse channel sensor configuration for surfaces. In this architecture, a few elements in the IRS remain active (without pilot symbols for channel estimation), and compressive sensing is employed to acquire the channel responses on all other passive elements. This knowledge can then be exploited in a deep learning-based solution to design the reflection matrices with no training overhead. For IRS systems with imperfect CSI, distributed reinforcement learning techniques are also considered for channel estimation in [316].

X. Extensions

A. THz Sensing, Imaging, and Localization

The unique THz spectral fingerprints of biological and chemical materials have been exploited in a variety of sensing and imaging applications [20], [25], [317]–[322], such as quality control, food safety, and security. As THz frequencies are close to the optical realm, the high-energy electromagnetic waves behave as photons, which often interact with other particles and matter. Such light-matter interactions with small particles (reflections, diffraction, and absorption) create unique electromagnetic signatures that can be exploited for THz sensing. THz signals can penetrate several materials and are strongly reflected by metals. They can also be used to analyze water dynamics (due to molecular coupling with hydrogen-bonded networks) and gas compositions (rotational spectroscopy). Following recent advancements in THz technology, however, and with the prospect of realizing THz capabilities in hand-held devices, THz sensing applications are expected to extend beyond the traditional industrial and pharmaceutical domains to reach everyday applications. In particular, sensing, imaging, and localization applications are expected to be piggybacked onto THz wireless communications [20].

Novel THz-specific signal processing techniques are required to enable efficient joint THz sensing and communications. The first stage of THz sensing is signal acquisition, which is usually achieved via THz time-domain spectroscopy (THz-TDS). THz-TDS can be realized in reflection mode or transmission mode, with the latter being more useful for sensing/imaging from a distance. Furthermore, reflection-based spectroscopy is more convenient in the context of joint communications and sensing. Following the signal acquisition, a variety of signal processing and machine learning techniques can be used to pre-process the received signals, extract characteristic features, and classify target materials into appropriate classes. Furthermore, the accuracy of sensing and imaging is greatly enhanced in the THz band due to the vastly wider available channel bandwidths, as well as the high directionality that comes with massive MIMO beamforming. Smart metasurfaces operating in the THz band are also capable of sensing environments [323]. Note that in the case of massive THz MIMO systems, carrier-based sensing/imaging can prove to be more efficient as multiple RF chains can be tuned to multiple frequencies, generating multiple responses over the THz spectrum. Compared to short pulses that cover the entire THz frequency range, carrier-based THz systems (frequency-domain spectroscopy) provide greater flexibility for choosing the carriers of interest for specific sensing applications. Note that only a few carefully-selected carriers can provide an efficient test for the existence of a specific molecule.

In the particular case of carrier-based THz-band wireless gas sensing (also known as electronic smelling [34]), the estimated channel responses can be correlated with the HITRAN database [90], so that a decision is made on the gaseous constituents of the medium. To illustrate that this sensing procedure can be seamlessly piggybacked over a communication system, we consider an UM-MIMO AoSA scenario and assume each SA to be tuned to a specific frequency in a symmetric manner at the transmitter and the receiver. Note that multiple SAs can still be tuned to the same frequency while assuming the channel to be orthogonal by design (following spatial tuning). Hence, the corresponding channel is diagonal, and the MIMO problem can be resolved into multiple single-input single-output problems. Each diagonal entry of $\mathbf{H}$ thus represents the channel response between a particular SA at the transmitter (tuned to a particular frequency) and its corresponding SA at the receiver side. The received vector can then be expressed as $\mathbf{h}_{m,n,m,n} = \mathbf{a}_r^H(\theta_r, \phi_r)G_r \bar{\alpha}_{m,n,m,n} G_t \mathbf{a}(\theta_t, \phi_t)$, where $\bar{\alpha}_{M,N_r,M,N_t} = \frac{c}{4\pi f d_{m,n,m,n}} e^{-j2\pi f d_{m,n,m,n}}$, with $\mathcal{K}^g(f_n)$ being the absorption coefficient of gas $g$ at
frequency $f_n$. By solving for $\mathbf{H}$, we identify both the gasses (or even specific isotopes of gases) that exist in the medium and their concentrations (by inspecting equation 3). The larger the AoSA size is, the more observations can be accumulated per channel use, and the faster the decision is made on the constituents of the medium.

Several methods can be used to solve for the absorption coefficients in $\mathbf{H}$, including optimal maximum likelihood detectors, variations of compressed sensing techniques, as well as machine learning algorithms. For instance, instead of comparing the exact values of channel measurements, we can set thresholds to check the presence or absence of specific spikes and build decision trees for classification \([324]\). Note that for sensing purposes, $\mathbf{x}$ can be assumed to be a random vector. However, in joint sensing and communications setups, the entries of $\mathbf{x}$ would belong to a specific constellation with a specific structure. This knowledge can be exploited to enhance the sensing performance further. In \([309]\), artificial intelligence is used in the context of IRS-assisted intercell mmWave communication, for sensing, programmable computing, and actuation facilities within each unit cell.

In addition to THz sensing and imaging, high-resolution localization capabilities in the THz band \([326] - [331]\) are key for THz communications, especially since the beams are narrow and mobile users are hard to track. Higher directional-ity and array compactness and larger bandwidths are all features that can be leveraged to enhance THz-based localization. Several well-established localization schemes can be extended to THz-band operations, such as ranging techniques based on time of arrival, time difference of arrival, angle of arrival, and received signal strength. High-frequency localization is mainly based on simultaneous localization and mapping (SLAM), which can leverage THz-generated high-resolution images of the environment to enhance the localization accuracy. THz network localization using multidimensional scaling (MDS) is also proposed in \([20]\). THz radar, further, promises to achieve millimeter accuracy. In the context of IRS systems, by acquiring the accurate position of user equipment, joint precoding at the base station and phase shifting at the IRS can guarantee accurate angle or delay estimation. In \([332]\), a leaky-wave antenna with a broadband transmitter is proposed for single-shot link discovery of neighboring nodes. Furthermore, learning-based techniques can leverage environment awareness for precoding at higher frequencies \([333]\). THz sensing, imaging, and localization applications can all be piggybacked onto THz wireless communication, or supported via dedicated resource allocation schemes \([20]\).

### B. Networking and Security

Having addressed several THz-specific signal processing techniques, it is worth noting that it is both the signal processing and networking problems that are significantly different in the THz band; both are linked to the underlying THz device architectures. Hence, THz-specific medium access (MAC) protocols \([334] - [340]\) need to be optimized jointly with physical layer signal processing schemes, under the constraints of state of the art THz devices. In particular, multiple access and networking paradigms for highly varying THz mobile environments are required. Examples of joint optimization schemes include the works in \([341]\), where energy harvesting THz nanonetworks are designed for controlling software-defined metamaterials, and in \([342]\), where joint THz power allocation and scheduling is optimized in mesh networks. Similarly, an on-demand multi-beam power allocation MAC protocol for THz MIMO networks is proposed in \([338]\). Synchronization of ultra-broadband THz signals, spectrum access and sharing, and neighbor discovery (given narrow beams) are open problems that require solutions on both the physical and network layers.

Security issues at the physical and network layers are also important to address. With higher propagation losses and increased directionality, THz communications are surely more secure than communication paradigms at lower frequencies. Nevertheless, this enhanced security is not perfect. Security and eavesdropping in THz links is first studied in \([343]\), where it is argued that security protocols should be designed on multiple levels, including hardware and the physical layer. Signal processing techniques for waveform design can be proposed for the latter. In particular, a scattering object can still be placed within the broadcast sector of a transmitting antenna, despite the increased spatial resolution, which would then scatter radiation towards a nearby eavesdropper. By perfectly characterizing the backscatter of the channel, such a security breach can be avoided. Furthermore, narrow beams can still cover a relatively large area around the receiver, which is a vulnerability that can be exploited for eavesdropping. In \([344]\), this vulnerability is mitigated by THz multipath propagation, at the expense of slightly reduced capacity.

The concept of covert THz communication, in which an adversary, when residing inside the beam sector, is prevented from knowing the occurrence of transmission, is also gaining attention. For instance, in \([345]\), covert THz communication is studied at the network level in the context of dense internet-of-things systems, where reflections and diffuse scattering from rough surfaces are exploited. Furthermore, in \([346]\), covertness is achieved by designing novel modulation schemes, such as distance-adaptive absorption peak hopping, in which frequency hopping is strategically selected at THz molecular absorption.
peaks. The covert distance is dictated by the transmit power and the SNR thresholds.

XI. Conclusions

In this paper, we present a first-of-its-kind tutorial on signal processing techniques for THz communications. We detail the THz channel characteristics and summarize recent literature on THz channel modeling attempts, performance analysis frameworks, and experimental testbeds. We further highlight problem formulations that extend classical signal processing for wireless communications techniques into the THz realm. In particular, we study THz-band modulation and waveform design, beamforming and precoding, channel estimation, channel coding, and data detection. We extend the discussion to cover the role of reflecting surfaces in the THz band, as well as THz sensing, imaging, and localization. We also shed light on THz-band networking and security issues. The techniques discussed in this paper will continue to evolve in the near-future, driven by advancements in THz transceiver design and system modeling.

References


S. K. Moorothy and Z. Guan, “FlyTera: Echo state learning for joint access and flight control in THz-enabled drone networks.”


**Biographies**

**Hadi Sarieddeen** (S’13–M’18) received his B.E. degree (summa cum laude) in computer and communications engineering from Notre Dame University-Louaize (NDU), Lebanon, in 2013, and his Ph.D. degree in electrical and computer engineering from the American University of Beirut (AUB), Beirut, Lebanon, in 2018. He is currently a postdoctoral research fellow at KAUST. His research interests are in the areas of communication theory and signal processing for wireless communication systems.

**Mohamed-Slim Alouini** (S’94–M’98–SM’03–F’09) was born in Tunis, Tunisia. He received his Ph.D. degree in Electrical Engineering from Caltech, CA, USA, in 1998. He served as a faculty member at the University of Minnesota, Minneapolis, MN, USA, then in Texas A&M University at Qatar, Doha, Qatar, before joining KAUST as a Professor of Electrical Engineering in 2009. His current research interests include the modeling, design, and performance analysis of wireless communication systems.

**Tareq Y. Al-Naffouri** (M’10–SM’18) Tareq Al-Naffouri received his B.S. degrees in mathematics and electrical engineering (with first honors) from KFUPM, Saudi Arabia, his M.S. degree in electrical engineering from Georgia Tech, Atlanta, in 1998, and his Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 2004. He is currently a Professor at the Electrical Engineering Department, KAUST. His research interests lie in the areas of sparse, adaptive, and statistical signal processing and their applications, localization, machine learning, and network information theory.